GEOFFREY K. PULLUM AND GERALD GAZDAR

NATURAL LANGUAGES AND
CONTEXT-FREE LANGUAGES*

0. INTRODUCTION

In his 1956 paper 'Three Models for the Description of Language' Noam
Chomsky posed an interesting open question: when we consider the
human languages purely as sets of strings of words (henceforth *string-
sets*), do they always fall within the class called *context-free languages*
(CFL's)? Chomsky declared that he did not know the answer to this
question, and turned to a very different set of questions concerning
relative elegance and economy of different types of description. Since
1956 various authors (Chomsky included) have attempted to provide
answers in the negative, and the negative answer is now the standardly
accepted one. We take up the question again in this paper, and show that
it is still open, as all the arguments for the negative answer that have
been provided in the literature are either empirically or formally incor-
rect.

The question of where the natural languages (as stringsets) are located
on the 'Chomsky hierarchy' (Type 0 or recursively enumerable languages,
Type 1 or context-sensitive languages, Type 2 or CFL's, Type 3 or finite
state languages (FSL's)) is an intrinsically interesting one. As Kimball
(1973, 26) remarks:

The (Chomsky hierarchy) represents the fact that regular languages [i.e. FSL's –
GKP/GJMG] are the simplest or least complex, the CF languages are next, and the CS are
the most complex of the phrase structure languages. In a certain sense, the problem faced
in the construction of a theory of universal grammar is to determine exactly how 'complex'
natural language is.

In the past fifteen years or so, concern with questions of 'generative
capacity' has declined somewhat, but, clearly, the mathematical proper-
ties of the object of study are of intrinsic interest regardless of whether
or not it is currently fashionable to stress such properties. Moreover,
there are certain very important applications of work on this question.
One area of application is the study of recognition algorithms and a
second is the study of language identification algorithms (cf. Gazdar
(in press) for discussion). The question we are concerned with is of
interest in connection with machine processing of natural language.
Compiler design for CFL's is a fairly well explored problem (see Aho
and Ullman 1972, 1973), but designing compilers for non-CFL's can be

grossly more difficult. Nor is this a concern that can be relegated to the field of computer programming; for those who take seriously the thesis of Fodor (1975), language acquisition for a human learner is nothing more or less than the construction of a program to compile the natural language into the human machine code (or whatever intermediate code is used for thinking).

Prescinding away from applications, however, the work of assessing the validity of claims about the alleged non-context free character of human languages is inherently rewarding in that in some cases it leads to the consideration of complex and descriptively interesting data (cf. especially the facts from Dutch and Mohawk discussed below). It also has the potential of being highly relevant to theoretical linguistics in a way that has not hitherto been noted. Suppose some piece of mathematical work were able to show of some theory – say, the arc pair grammar of Johnson and Postal (1980) or the 'government-binding' theory of Chomsky (1981) – that it necessarily only defined grammars for CFL's.[1] If it were generally believed that the natural languages had already been shown not be a subset of the CFL's, the theory would immediately be charged with inadequacy, on the grounds that for some natural languages it would not provide for even a weakly adequate grammar (i.e. one that generated the right set of strings). But if the general belief about natural languages were incorrect, the theory (in principle, perhaps a correct theory) would have been unwarrantedly dismissed. As an insurance against such easily conceivable situations arising, we need to be sure that questions of weak generative capacity have not been incorrectly answered and prematurely closed.

We shall discuss first, in Section 1, the general background to the issue, and some of the prevailing misconceptions embodied in the linguistic literature. We then consider the only five published arguments for the non-context-freeness of natural languages: one involves comparative clauses in English (Section 2); one applies the pumping lemma directly to English (Section 3); one hinges on constructions with *respectively* (Section 4); one is based on facts of Dutch 'verb raising' sentences (Section 5); and one depends on aspects of noun incorporation in Mohawk (Section 6). In Section 7 we offer our conclusions.

1. FOLKLORE

Introductory textbooks on generative grammar often mention the alleged 'inadequacy' of context-free phrase structure grammar (CF-PSG) for the description of natural language. But even the best of them offer only non

sequiturs as backing for such denigrations. A typical example from what is widely acknowledged to be one of the best textbooks of its type is the following piece of reasoning from Akamajian and Heny (1975, p. 86), which follows their introduction of a phrase structure rule for 'AUX' which fails to allow for a phrase structure description of auxiliary-initial interrogatives in a non-redundant way:

Since there seems to be no way of using such PS rules to represent an obviously significant generalization about one language, namely, English, we can be sure that phrase structure grammars cannot possibly represent *all* the significant aspects of language structure. We must introduce a new kind of rule that will permit us to do so.

There are several non sequiturs here, the central one being that the fact that "there seems to be no way of using such PS rules" for some task does not license the inference that no successful phrase structure (CF-PSG) account could possibly be devised.[2]

Culicover (1976, 50) argues very similarly, also on the basis of the English auxiliary system, claiming that it is "impossible to capture the observed generalizations in purely phrase structure terms". And, as pointed out by Hurford (1980, pp. 135ff) in a perceptive review of the two texts just cited, Culicover makes other such claims elsewhere in his book. After a discussion of *wh*-movement constructions he concludes that "the phrase structure analysis will not be sufficient to generate English" (p. 28), and three pages later he diagrams an abstract configuration (somewhat reminiscent of subject-verb concord) that he claims "cannot be described by phrase structure rules alone" (p. 31). Culicover is quite wrong in his claims about phrase structure grammar; simple CF-PSG's can be constructed (and are constructed by Hurford) to describe the phenomena cited.

The belief that CF-PSG is inadequate to cope with long-distance dependencies, syntactic concord, and so on, is well entrenched. Grinder and Elgin (1973, p. 57) go so far as to claim that even verb agreement in simple sentences "demonstrates clearly the inadequacy of ... the CF-PSG". They exhibit some simple examples, and then assert that

any set of CF-PSG rules that generate (correctly) the sequences [*The girl kisses the boy*] and [*The girls kiss the boy*] will also generate (incorrectly) the sequences [**The girl kiss the boy*] and [**The girls kisses the boy*]. The grammatical phenomenon of Subject-Predicate Agreement is sufficient to guarantee the accuracy of [the statement that] English is a not a CF-PSG language (p. 59).

The phenomenon guarantees no such thing, of course.[3]

Nor is the character of the problem changed when agreement-rule effects are exhibited across unbounded distances in strings. Yet Bach (1974, p. 77) states that

to describe the facts of English number agreement is literally impossible using a simple agreement rule of the type given in a PSG, since we cannot guarantee that the noun phrase that determines the agreement will precede (or even be immediately adjacent to) the present-tense verb.

And Bresnan (1978, p. 38) makes the following claim:

In contrast to the local type of number agreement ..., the distant type of agreement ... cannot be adequately described even by context-sensitive phrase structure rules, for the possible context is not correctly describable as a finite string of phrases.

Bresnan is referring to the interaction of *wh*-extraction and number agreement in examples like *Which problems did your professor say she thought were unsolvable?* It is perhaps worth demonstrating that far from being beyond the capacities of even context-sensitive phrase structure grammar, the phenomenon Bach and Bresnan are referring to can be described by a *finite state* grammar. Here is a simple finite state grammar with only three nonterminal symbols which generates an infinite subset of English including the example just cited.[4]

> S → *Which problem did your professor say* T
> S → *Which problems did your professor say* U
> T → *she thought* T | *you thought* T | *was unsolvable?*
> U → *she thought* U | *you thought* U | *were unsolvable?*

This is not a proposed fragment of the grammar of English, of course; it is presented merely as a simple proof that an infinite language with long-distance dependencies and syntactic concord over unbounded domains can be an FSL.[5]

The views we have been discussing seem to have become a kind of folklore within generative grammar, frequently repeated for the benefit of the young though not provided with any scientific backing. We find the familiar claims about the 'inadequacy' of CF-PSG reiterated or simply assumed in linguistic works of all kinds. We find introductory texts, beside the ones already cited, saying things like, "Phrase-structure rules are not sufficient to account for all syntactic knowledge" (Fromkin and Rodman, 1979, pp. 234–235) and "We must assume that any phrase-structure grammar of a natural language will need to make extensive use of context-sensitive rules" (Allerton, 1980, p. 82). We find technical papers in syntax making similar claims, as when Selkirk (1977, p. 285) takes it to be "the single most important contribution to the development of linguistic theory in this century" that "the inadequacy of phrase structure grammars as a model of linguistic structure" has been demonstrated.[6] In discussions of parsing we find Winograd (1972, p. 42) asserting that

"context-free parsers...cannot handle the full complexity of natural language", and Langendoen (1975, p. 536n) claiming that "the generative capacity...of CFPS grammars is too small to be of linguistic interest". And in Pinker's review of the formal literature on learnability we find the assumption that "natural languages are not context-free" being accepted without remark as if it were quite uncontroversial (Pinker, 1979, p. 223).

The introductory texts and similar expository works in the field of generative grammar offer nothing that could be taken as a serious argument for the claim that natural languages are not CFL's, and yet the general assumption in the literature is nevertheless that this claim is true. It is reasonable, therefore, to assume that everyone is tacitly relying on more technical arguments available elsewhere, which are presumably thought too complex for pedagogical works but nonetheless trustworthy. What form would a trustworthy argument take? A natural language (and indeed, any language) will have an infinite number of grammars that generate it, and one clearly cannot search through an infinite set of grammars to see if one of them is a CF-PSG. Hence a demonstration by enumeration of candidate grammars in order to show that they fail (the standard expository technique) is out of the question. We must utilize the various theorems about CFL's that have been proved in purely mathematical work on properties of stringsets.

One theorem about CFL's is the so-called 'pumping lemma'. Only once to our knowledge has anyone attempted to apply the pumping lemma directly to a natural language, and we discuss the (unsuccessful) result below in Section 3.

More useful than the pumping lemma are two closure properties of the CFL's: they are closed under intersection with FSL's, and closed under homomorphisms. The former property means that if one can define an FSL F such that $L \cap F = L'$ and L' is not a CFL, then neither is L. The latter property implies that if one can construct a homomorphism (a sequence-preserving symbol-by-symbol mapping) that converts L to L' and L' is not a CFL, then neither is L.

To use these properties to prove theorems about natural languages it is necessary to ensure that L' is in each case transparently non-context-free, ideally a language that can directly be shown not to be a CFL by application of the pumping lemma. Most published arguments have depended on setting some natural language as L and (implicitly in some cases) letting L' be an instance of what we shall call an *xx language*. The canonical instances of *xx* languages are $\{xx \mid x \in L((a + b)^*)\}$ and $\{xx \mid x \in L((a + b)^+)\}$. We shall call a language an *xx* language just in case it can be mapped into one or other of these under the operations of

homomorphism and intersection with FSL's. All xx languages in this
sense are non-context-free. The proof that the canonical xx languages
are not CFL's is straightforward (see, e.g., Aho and Ullmann (1972, p.
198)), and the non-context-freeness of xx languages in general im-
mediately follows from the closure properties of CFL's mentioned
above. Therefore, if a natural language can be shown to be an xx
language, that natural language will not be a CFL. Sections 4, 5, and 6
analyze all the published attempts to obtain such a result that we are
aware of. First, however, we must dispose of two unsuccessful
arguments that do not have this form, in Sections 2 and 3.

2. COMPARATIVES

Chomsky (1963, pp. 378–9) based an argument on what we shall refer to
as an xy language, i.e. one whose grammar requires *non*identity
between substrings x and y in its sentences.

... it should be observed that a language is also beyond the weak generative capacity of
context-free grammars ... if it contains an infinite set of phrases $x_1, x_2 \ldots$, and sentences
of the form $\alpha x_i \beta x_j \gamma$ if and only if i is *distinct* from $j \ldots$. Thus in the comparative
construction we can have such sentences as *That one is wider than this one is DEEP* (with
heavy stress on *deep*), but not **That one is wider than this one is WIDE* – the latter is
replaced obligatorily by *That one is wider than this one is*. Thus in these constructions,
characteristically, a repeated element is deleted and a nonrepeated element receives heavy
stress. We find an unbounded system of this sort when noun phrases are involved, as in the
case of such comparatives as *John is more successful as a painter than Bill is as a
SCULPTOR*, but not **John is more successful as a painter than Bill is as a PAINTER ...*,
these constructions show that natural languages are beyond the range of the theory of
context-free grammars or PDS automata, irrespective of any consideration involving
strong generative capacity.

Chomsky is claiming two things. First, that English contains a set of
sentences of the form $\alpha x \beta y \gamma$ where β is arbitrary central material (such
as (*-er*) *than this one is*) and the phrases x and y have to meet the
condition that they are *not* identical.[7] And second, that a set of strings
meeting these conditions (an 'xy language') cannot be enumerated by a
CFG, so English cannot either.

We consider the empirical premise first. Chomsky, as pointed out by
Bresnan (1976, pp. 373–4), no longer regards sentences such as (1) below
as ungrammatical.

(1)   That one is wider than this one is WIDE.

In fact, in the course of arguing against Bresnan's analysis of com-
parative clauses, he cites the example in (2):

(2)   What is more, this desk is higher than that one is HIGH.
       (Chomsky, 1977, p. 122)

We are entirely persuaded by Chomsky's argument that this sentence is grammatical. We presume that the same should be said of (1), and that Chomsky's original judgement on it (1963, p. 378) was wrong.

There is a second problem with the factual basis of the argument. Green (1971, p. 560) observes that sentences like (3) seem very peculiar (ungrammatical, in her view).

(3)   John is as fat as Bill is obese.

This is an equative; but the same kind of oddness is felt in (4).

(4)   John is fatter than Bill is obese.

The linguistic strangeness of these sentences seems to be due to the fact that the two adjectives employed are synonymous, so that in virtually any context it would have been a good deal simpler and less puzzling to have used the elliptical variants (5), (6).

(5)   John is as fat as Bill is.

(6)   John is fatter than Bill is.

But exactly the same can be said regarding (1) for this has an elliptical variant as well, as Chomsky stresses:

(1)   That one is wider than this one is WIDE.

(7)   That one is wider than this one is.

The only circumstance in which (1) would not share the oddness of examples like (3) and (4) would be one in which (7) did not sufficiently emphasize that the adjective in the *than*-clause was *wide* as opposed to some alternative with which the speaker wanted to contrast it. In other contexts, whatever semantic or pragmatic explanation covers (3) and (4) – and there must be one – will also cover (1). Thus whatever residual linguistic oddness inheres in Chomsky's crucial examples has an independent explanation unrelated to questions of superficial nonidentity of adjectives. The syntactic nonidentity condition is not only unnecessary (given that (1) is grammatical), it would also have been insufficient in failing to eliminate (3) and (4).

We now turn to the formal side of the argument, which is so flawed that any remaining uncertainty about the data discussed above is an irrelevance.

Chomsky does not actually give a demonstration that a homomor-
phism or an intersection with a finite state language could connect
English in the appropriate way to an xy language. We can ignore this
because there is a far more serious formal flaw in his argument.
Although Chomsky's crucial premise is that xy languages are inherently
non-context-free, he neither supplies nor alludes to a proof of this. Nor
could he, in fact, for it is false. There are infinitely many context-free xy
languages. To illustrate, we shall take an example based on the
specifications given by Chomsky above. We assume a terminal vocabu-
lary $\{a, b, \alpha, \beta, \gamma\}$, and consider the language denoted by (8).

(8) $\{\alpha x \beta y \gamma \mid x, y \quad L((a + b)^*) \wedge x \neq y\}$

This is a CFL. Rather than use the pumping lemma to prove it, we shall
simply exhibit a CF-PSG that generates the language.[8]

(9)   (a) $S \rightarrow \alpha S' \gamma \mid \alpha S'' \gamma$
        (b) $S' \rightarrow CS'C \mid D\beta \mid \beta D$
        (c) $S'' \rightarrow AB' \mid BA'$
        (d) $A \rightarrow CAC \mid a(D)\beta$
        (e) $B \rightarrow CBC \mid b(D)\beta$
        (f) $A' \rightarrow a(D)$
        (g) $B' \rightarrow b(D)$
        (h) $C \rightarrow a \mid b$
        (i) $D \rightarrow C(D)$

Since it is rather hard to see that the language described in (8) is a CFL
even when (9) is provided, a few words of explanation are in order. The
grammar in (9) develops terminal strings in three main ways. Rule (9a)
starts things off by placing α at the beginning and γ at the end of the
string. Recursion through S' using the first alternative in (9b) yields an S'
flanked by equal numbers of a's and/or b's. To stop the recursion, use
the second alternative in (9b), which places a D and the center marker β
in that order. D can dominate any number greater than one of a's and/or
b's. Thus there will be at least one more character before the β than after it,
so the x string (between α and β) will differ in length from the y string
(between β and γ). The third alternative in (9b) handles the converse case
(more characters in y than in x) analogously. This leaves the problem of
generating the strings where x is the same length as y but x and y are
different. For x to differ from y other than in length, there must be some j
such that at the jth character, x has an a and y has a b or conversely. To
ensure the former, use the second alternative of (9a), and then expand S'' by
means of the first half of (9c) to derive AB'. Recursion i times on (9d) leads
to an A with i instances of a or b characters on each side. To stop this

recursion, use the second alternative in (9d) to introduce an a followed by an optional further sequence of a or b characters (dominated by D) and a β. Expand B' via (9g), as b followed by an optional string of a's and/or b's. The specified a introduced by (9d) will be preceded by exactly i characters from the set $\{a, b\}$, and exactly i characters from this set will separate β from the specified b that (9g) supplies. Thus x will differ from y at the $i + 1$th symbol: x will have an a where y has a b. The converse is also provided for in an obvious way via recursion on B (using rules (9e) and (9f)). These hints should suffice to enable the reader to check with paper and pencil that the grammar in (9) guarantees dissimilarity between the $\alpha \ldots \beta$ and $\beta \ldots \gamma$ strings and thus generates the language described in (8).

This proves that the formal premises for Chomsky's argument are as invalid as the empirical ones, and the entire argument is unsalvageable.[9]

3. Pi

In a recent book concerned with applying logical theory to the social sciences, Elster (1978) touches on the question we are considering in this paper, and offers an original argument purporting to show that English is not a CFL. The argument involves applying the pumping lemma directly to a set of English sentences about the decimal expansion of pi. Since the argument is fairly concise, we reproduce it in full below*.

Bar-Hillel et al. have proved the following theorem:

Let A be a context-free grammar and L(A) the set of sentences generated by A. Then there exists a number p such that for any z in L(A) longer than p, there exist x, u, w, v, y such that

    (i) $z = xuwvy$,

    (ii) $xu^m wv^m y \in L(A)$ *for all m*;

    (iii)   *u and v are not both empty.*

More briefly the theorem states that in a language generated by a context-free grammar any sufficiently long sentence can be extended by indefinite repetition of *two* subparts without violation of grammaticality, just as we know that in languages generated by a finite state grammar any sufficiently long sentence can be extended by indefinite repetition of one subpart. In order to prove the inadequacy of context-free grammars for natural languages, we must come up with a counterexample of a grammatical sentence that cannot be extended in this way. The counter-examples of the form 'If^k it blows, (then it rains)^{k}' will not do for this purpose, for these sentences *can* be indefinitely extended by repetition of the two exponential blocks. Now let us look at the following sentences:

B_1: The first two million numbers in the decimal expansion of π are $a_1 a_2 \ldots a_{2000000}$.

B_2: The first two million numbers in the decimal expansion of π are $a_1 a_2 \ldots a_{2000000000000}$.

    $\vdots$

B_k: The first two (million)^k numbers in the decimal expansions of π are $a_1 a_2 \ldots a_{2 \cdot 10^{6k}}$.

    $\vdots$

Given a context-free grammar and the number p of the above theorem, it is easy to find a k

such that B_k is longer than p. Such a B_k must then permit a double extension as stated in the theorem. It is not difficult to see that these extensions must occur within the blocks 'million^k' and '$a_1 \ldots a_{2.10^{6k}}$' if the result is not to be obviously ungrammatical. Suppose, then, that u in the theorem has the form 'million^q' and v the form '$a_r \ldots a_t$'. Choosing $m = 2$ in the theorem, it then states that the following sentence must be a grammatical one if we assume that the context-free grammar in question is adequate for English.

C: The two million^{k+q} first numbers in the decimal expansion of π are $a_1 \ldots a_t a_r \ldots a_t \ldots a_{2.10^{6k}}$

I shall show that C is not a grammatical sentence in English. If C is grammatical, then the number 'two million^{k+q}' must be the same as the number '$2.10^{6k} + t - r + 1$', i.e. the same as the number of numbers in the decimal expansion. Note that this is a requirement not of mathematics, but of linguistics, just as the lack of grammaticality of the sentence,

D: the two largest animals in the zoo are a mouse,

is a matter of linguistics, and not of mathematics (or of zoology). But as $q \geqslant 1$ and $t - r + 1 \leqslant 2.10^{6k}$, we have

$$2.10^{6k} + t - r + 1 \leqslant 4.10^{6k} < 2.10^{6k+1} < 2.10^{6(k+q)},$$

whereas the grammaticality of C would require $2.10^{6k} + t - r + 1 = 2.10^{6(k+q)}$.

The flaw in this argument is not hard to detect. Elster is assuming that the following principle is a rule of English grammar:

(10)   In the construction 'The W_1 are W_2', the number of entities listed in W_2 must correspond to the number named in W_1.

His only evidence for this principle is the ungrammaticality of (11a).

(11)(a) * The two largest animals in the zoo are a mouse.
     (b)   The two largest animals in the zoo are a mouse and a gerbil.

But the ungrammaticality of (11a) does not require his principle as an explanation, since it already follows from one of the most familiar, and least controversial, rules of English grammar, namely the rule that requires predicates to agree with their subjects in number. This rule accounts for the contrast between (11a) and (11b), as well as for the patterns of grammaticality shown in (12) and (13):

(12)(a)   Mickey is a mouse.
     (b) *Mickey are mice.
     (c) *Mickey and Minnie is a mouse.
     (d)   Mickey and Minnie are mice.

(13)(a)   The largest animal in the zoo $\left\{ \begin{matrix} \text{is} \\ *\text{are} \end{matrix} \right\}$ Mickey.

     (b) *The largest animal in the zoo $\left\{ \begin{matrix} \text{is} \\ \text{are} \end{matrix} \right\}$ Mickey and Minnie.

(c) *The two largest animals in the zoo $\begin{Bmatrix} \text{is} \\ \text{are} \end{Bmatrix}$ Mickey.

(d) The two largest animals in the zoo $\begin{Bmatrix} \text{*is} \\ \text{are} \end{Bmatrix}$ Mickey and Minnie.

However, crucially, the following strings *are* grammatical sentences of English:

(14) (a) The two largest animals in the zoo are Mickey, Minnie and Donald.
   (b) The three largest animals in the zoo are Mickey and Minnie.

Such sentences are necessarily false (or, perhaps, necessarily truth-valueless), and consequently infelicitous qua utterances. Elster's mistake is to confuse grammaticality with felicity. His argument based on pi has no bearing on English or any other natural language. There would be a few more cases to deal with grammatically in a language with morphological distinctions between singular, dual, and trial, but larger numbers govern plural concord in such languages, and our point could be made in exactly the same way. Elster's assumption that the sentences in (14) are ungrammatical is wrong in the same way that it would be wrong to try to achieve a syntactic account of what is peculiar about the examples in (15).

(15) (a) Here are six random integers: 3, 17, 8, 9, 41.
   (b) Our three main weapons are fear, surprise, ruthless efficiency, and a fanatical devotion to the Pope.
   (c) There'll be Bob, Carol, Ted, Alice, Bruce, Martha, Mike, and that's eight including the two of us.

In short, Elster's argument depends on a confusion between grammar and arithmetic.

4. RESPECTIVELY

Bar-Hillel and Shamir (1960, p. 96) present the earliest argument that English is an *xx* language and thus not a CFL. They suggest that a string of the form

> John, Mary, David, ... are a widower, a widow, a widower, ..., respectively

is a grammatical sentence if and only if the three dots are replaced by a string of any length of proper names, with repetitions and commas between the names, and the three dashes by a string of equally many

phrases of one of the forms 'a widower' or 'a widow' and such [that] whenever the nth proper name is male (or female), the nth phrase is 'a widower' (or 'a widow').

They then say that they regard this as "almost conclusive" proof that English is not a finite state language, and that it "shows also the inadequacy of the simple phrase structure grammars".

Daly (1974, pp. 57–60) gives a detailed critique of their argument as it stands, pointing out that what it asserts about English is quite untrue, and that formally its authors have not even begun to make a case that English is not CF. But we shall not review Daly's discussion, for what is important is not the failure of Bar-Hillel and Shamir to make an argument but the possibility that a valid argument might be based on *respectively*-constructions.

Langendoen (1977, pp. 4–5) has attempted to reconstruct the *respectively* argument with different examples and a secure formal basis. He first defines the language that we exhibit as the regular expression in (16);

(16)   **(the woman + the men)$^+$ and (the woman + the men) (smokes + drink)$^+$ and (smokes + drink) respectively.**

The expression $(a + b)^+$ denotes the set of all non-null strings consisting of a's and b's, so (16) contains an infinite set of strings like *the woman the men the men and the men drink drink smokes smokes and smokes respectively*, where no matching up between the singular and plural noun phrases and verbs need be present. This simply definable language is an FSL. Langendoen then considers the intersection of (16) with English. This, he claims, will be the language defined in (17):

(17) $\{xx'$ *respectively* $\mid x \in L(($**the woman + the men**$)^+$ **and (the woman + the men**$))$ and x' is the corresponding string that has *smokes* in place of *the woman* and *drink* in place of *the men*$\}$

Langendoen notes, correctly, that (17) is not a CFL; but since the intersection of a CFL and an FSL is always a CFL, and since (17) is (putatively) the intersection of (16) and English, English cannot be a CFL.

This would be a sound argument if its premises were all true, but we dispute one of the premises, namely the characterization of English that is assumed. The argument crucially turns on the grammatical status of the following sentences, which we do not mark with our own judgements but leave to the reader to assess or test out with informants before continuing.

(18) (a) The woman and the men smokes and drinks respectively.
  (b) The woman and the men smokes and drink respectively.
  (c) The woman and the men smoke and drinks respectively.
  (d) The woman and the men smoke and drink respectively.

Our judgements are as follows. All of (18 a–c) we find ungrammatical. While (18d) is of uncertain status, it is the only one we could conceivably accept (assuming that the meaning intended for each of these strings is 'The woman smokes and the men drink'). But what the *respectively* argument uses as a premise is a characterization of English under which (18b) alone is grammatical, the others all being ungrammatical. This does not seem correct to us (and nor does it seem correct to Geoffrey Sampson, who makes the same point we are making here, on the basis of a different example, in Sampson (1975, pp. 205–206)).

Even more tellingly, consider the data in (19).

(19) (a) The man and the woman smokes and drinks respectively.
  (b) The man and the woman smoke and drink respectively.

We find we can tolerate (19b), but we reject (19a) completely. This is the exact converse of the grammaticality assignments assumed in Langendoen's characterization of English *respectively* constructions on subject-verb pairs.

We also find, more generally, that it is not true, as every version of the *respectively* argument seems to assume, that there is some numerical matching between syntactic elements in every grammatical sentence containing *respectively*.[10] Simple expositions of the *respectively* construction generally point to strings like (20) and call them ungrammatical.

(20)   Ira, Walt, and Louise have been dating Frank, Edith, Cedric, and Bruce, respectively.

Such sentences (of which more below) are indeed odd; but that is not because of some one-to-one mapping requirement on syntactic constituents. Suppose that someone has just been attempting to pick out, on a mass photo of the LSA membership, all those who reside in New York, and the last two faces they point to are those of Jerry Sadock and Arnold Zwicky. One might utter (21):

(21)   The people you have indicated are all New York residents except for the last two. They live in Chicago and Columbus respectively.

In this case, the nonlinguistic context picks out a unique ordered set of people with whom Chicago and Columbus can be appropriately asso-

ciated; but the sentence that contains *respectively* has an odd number of noun phrases in it, so pairing of constituents certainly cannot be part of the correct characterization of *respectively* sentences.[11]

If the oddity of (21) does not arise from the odd number of noun phrases it contains, then, what causes it? De Cornulier (1973) provides what we think is the right answer (see also, Hinton 1978). The way Langendoen's or any other version of the *respectively* argument treats the item *respectively* implies that it is without inherent meaning. While all other adverbs in the language will be lexical items with their own semantic contribution to make to the meaning of sentences in which they occur, *respectively* will be a syncategorematic item, introduced as a property of the syntactic construction in which it occurs (the way the past participle-forming morpheme-*en* is often treated in transformational analyses of the passive construction). De Cornulier suggests that it has a straightforward meaning, albeit a metalinguistic one. It can be roughly paraphrased as 'each in the order in which I cite them'. (One might revise this to 'each in the order indicated or implied', given cases like (21).) He points out that (22a) and (22b) – which he gives in French, though nothing changes under our translation into English – seem odd to the same degree and intuitively for the same reason:

(22) (a) Paul, Mary and Jack are respectively (male) cousin, uncle, and aunt to Robert.
     (b) Paul, Mary and Jack are, in the order in which I cite them, (male) cousin, uncle, and aunt to Robert.

But it would be absurd to attempt a syntactic account of why (22b) is odd. And as soon as an appropriate semantic or pragmatic account is set up, it will generalize to (22a), provided that *respectively* is assigned a meaning similar to the one de Cornulier suggests for it.[12]

Since *respectively* constructions have figured in arguments that phrase structure grammar must be abandoned in favour of transformational grammar, it is worth observing that no observationally adequate syntactic description of these constructions was ever provided in the transformational literature. This is presumably a rather surprising and disturbing fact if the construction in question really was one of the main reasons for the complete abandonment of phrase structure grammar in favour of the less constrained transformational model of syntactic theory.

Summarizing, we find that in the case of the *respectively* argument, the crucial data purporting to show English not to be CF are firstly mistaken at the level of observational adequacy, and secondly, illustrative only of

semantic or pragmatic points insofar as they illustrate anything. We consider the *respectively* argument, therefore, to be completely without force.

5. DUTCH

An ingenious attempt to show that Dutch is not context-free has been made by Huybregts (1976). This article was only published in working-paper form and may not represent the author's current views, so rather than address ourselves closely to the text of Huybregts' paper, we shall address the facts, and show that no argument of the sort Huybregts advances can in fact succeed.

Although the verb is final in the verb phrase of a simple subordinate clause in Dutch, when transitive infinitival *VP*'s are nested the verbs do not appear in the order that would be predicted simply by nesting one *VP* inside another as in (23).

(23)
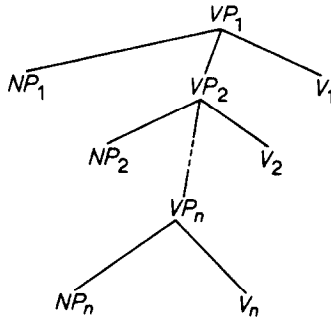
Instead, the string will generally look like this:

(24) $NP_1 \quad NP_2 \ldots NP_n \quad V_1 \quad V_2 \ldots V_n.$

Here $V_1 \ldots V_{n-1}$ are taken to be verbs that select a direct object *NP* and a complement *VP* (cf. *make, let*, or *help* in English), and V_n is some transitive verb. For all i (where $1 \le i \le n$), NP_i is the direct object of V_i, and is present because V_i is subcategorized to require it. Thus we have the beginnings of an argument of the same sort as the *respectively* argument, it might be claimed: Dutch has an infinite subset with an unbounded cross-serial dependency of the type $a_1 a_2 \ldots a_n b_1 b_2 \ldots b_n$, which can be mapped by a homomorphism into an *xx* language, and thus is not context-free. However, the actual situation is rather different, as we shall show, and ultimately no such argument can be made.

Consider some concrete examples. Example (25) is a grammatical subordinate clause in Dutch.

(25)   dat Jan [Marie Pieter Arabisch laat zien schrijven]
         that Jan Marie Pieter Arabic   let see   write
         'that Jan let Marie see Pieter write Arabic'

The bracketed portion contains three verbs (the first finite, the others
infinitival) and three *NP*'s. The first *NP* is, in some sense, the direct
object of the first *V* and the subject of the second; the second *NP* is, in
the same sense, the object of the second *V* and the subject of the third;
and the third *NP* is the object of the third *V*.

Not all verbs are transitive, of course. When selected verbs in the
above example are replaced by verbs of other valencies, the number of
NP's required changes. If *schrijven* 'write' is replaced by *liegen* 'lie',
(25) becomes the ungrammatical (26):

(26)   *dat Jan [Marie Pieter Arabisch laat zien liegen]
         that Jan   Marie Pieter Arabic   let see lie
         '*that Jan let Marie see Pieter lie Arabic'

Similarly, ungrammaticality results if the number of *NP*'s is altered, say
by adding *Hans*:

(27)   *dat Jan [Marie Pieter Arabisch Hans laat zien schrijven]
         that Jan Marie Pieter Arabic   Hans let see   write
         '*that Jan let Marie see Pieter write Arabic Hans'

However, additional verbs can be inserted if they are *intransitive VP*-
complement-taking verbs:

(28)   dat   Jan [Marie Pieter Arabisch wil laten zien schrijven]
         that Jan Marie Pieter Arabic   will let   see write
         'that Jan will let Marie see Pieter write Arabic'

Thus the number of *NP*'s that can be permitted in such strings is a
function of the number of *V*'s and the valencies of those *V*'s.

However, this does not mean that the stringset of Dutch is non-
context-free. Even when we keep in mind that the first verb in each
example is finite and all the others are in their infinitive forms, and that
the final verb has to be one that does *not* take a *VP* complement, it is
trivially easy to write a CF-PSG to assign the right number of *NP*'s to
go with the number of *V*'s selected.

Take first the case where only ordinary *NP*'s are involved (i.e.
excluding verbs that take *PP* complements, verbs like *regenen* that
demand a particular *NP* (*het* 'it') as subject, and so on). What we have
to do is to ensure that each of the relevant class of *VP*'s contains some

number n of NP's, a finite VP-complement-taking verb, a string x of nonfinite VP-complement-taking verbs, and a final transitive or intransitive verb, such that x contains $n - 2$ transitive verbs if the finite verb and the last verb are both transitive, $n - 1$ if only one of the finite verb and the last verb is transitive, and n transitive verbs otherwise. This will guarantee that each transitive verb has an NP to be its direct object, and that no intransitive verbs will get objects. The following grammar does it. (We use an arbitrary initial symbol A. Note that the grammar generates subordinate clause verb phrases, not sentences, and that for simplicity all NP's are taken to be simple names.)

(29) (a) Syntactic rules

$A \rightarrow BCD \mid CE$

$C \rightarrow BCF \mid CG \mid BH \mid I$

(b) Lexicon

B: *Marie, Pieter*, other personal names

D: *schrijven*, other transitive infinitives

E: *liegen*, other intransitive verbs

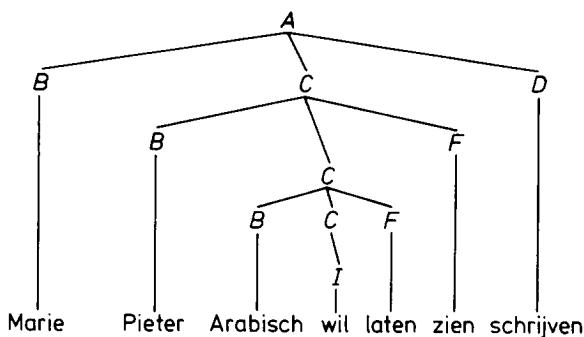F: *laten*, other transitive VP-complement-taking infinitives

G: *willen*, other intransitive VP-complement-taking infinitives

H: *laat*, other finite transitive VP-complement-taking verbs

I: *wil*, other finite intransitive VP-complement-taking verbs

The tree this grammar assigns to the VP of example (28) is shown in (30).

(30)

The grammar provides for intransitive verbs to occur interspersed with the transitive ones (*laten, zien*) shown in this example. Using *willen* yields for the most part very peculiar results for semantic reasons. Category I should be assumed to be augmented with such verbs as *scheen te* 'seemed to' (with *te* treated, to avoid complications, as part of the verb), and category G should likewise be assumed to contain

schijnen te 'to seem to' etc. These details do not affect the point under discussion here. For a summary of the behavior of various verbs in this construction, see Zaenen (1979).

What this grammar does not guarantee, quite deliberately, is that all the sentences it provides for will mean something sensible. For example, sentences like the following will be generated:

> (31)   (dat   Jan) Marie Arabisch Pieter wil laten zien schrijven
>   (that   Jan) Marie Arabic   Pieter will let   see write
>   '(that Jan) will let Marie see Arabic write Pieter'

This is, of course, a 'selection restriction' violation. The non-syntactic status of selection restrictions is surely quite uncontroversial by now. It is not for the syntax to rule out examples of this sort, for the above example is perfect under the assumption that there is a language or writing system called 'Pieter' and a person named 'Arabisch' has learned to write it.

When we turn to examples involving verbs that take complements other than *NP* (like *schrijven*) or nothing (like *liegen*), things are only slightly more difficult. Suppose there are n classes of verbs in Dutch, V_1 thru V_n, each taking different combinations of *NP*'s *PP*'s, or whatever as their subcategorized complements. Let us use the notation X_i for whatever string of phrases V_i is subcategorized to take. All we need is a set of n categories C_1 thru C_n, and for each verb class V_i we can provide a set of rules as shown in (32), the lexicon apart from the entry for V_i being shown in (29b).

> (32) $A \rightarrow C_i V_i$
> $C_i \rightarrow BC_iF \mid C_iG \mid BX_iJ \mid X_iI$
> $J \rightarrow H(E) \mid I(E)F$
> $E \rightarrow G(E)$

An example of how this grammar would generate a ditransitive case with *geven* 'give' is provided in (33). We assume that *geven* belongs to a class called V_3, so the rule schema '$C_i \rightarrow B X_i J$' in (32) is instantiated here by a rule '$C_3 \rightarrow B B B J$', allowing the appearance of *geven* to correlate with the appearance of a pair of extra *NP*'s:

> (33) (a) (dat   Jan) Marie Pieter Hans Fido moet laten zien geven
>   that Jan Marie Pieter Hans Fido must let   see give
>   '(that Jan) must let Marie see Pieter give Hans Fido'

(b)

Marie Pieter Hans Fido moet laten zien geven
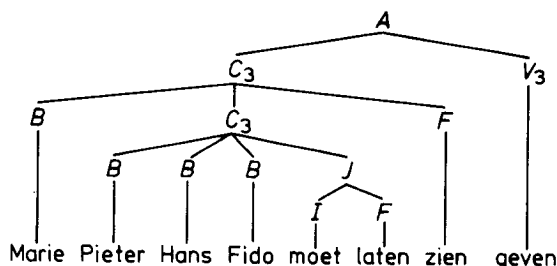
Allowing for a subject that is (let us assume) syntactically selected, like the *het* 'it' that goes with *regenen* 'rain', is also very easy. The above grammar makes X_i the final part of the preverbal half of the *VP* and V_i the final verb. By using the same mechanism we could make X_j be *het* where V_j is *regenen*. But this would give a word order that happens not to occur. It seems that *het* always assumes a special clitic position at the beginning of the *VP*. This is even easier to generate. Adding the rule '$A \rightarrow het\ C\ regenen$' to the grammar in (29a) is all that is necessary.

We must stress at this point that the foregoing examples of grammars and trees are not proposals concerning the grammar of Dutch or its constituent structure. The reader will note, for example, that the trees we have drawn are quite unsuitable for compositional semantic interpretation: semantic rules for assigning correct interpretations to Dutch sentences, constituent by constituent from the bottom up, cannot (as far as we can see) be provided on a rule-by-rule basis for the grammars given above. We do not intend to imply that they could. Although we have hypotheses in mind concerning the correct way to describe Dutch, this is not the place to develop them.[13] Our grammars are simply aids to showing, rather informally, that the facts we find in Dutch subordinate verb phrases are not incompatible with Dutch being a CFL.

Even syntactically, many details of Dutch are being glossed over, and many complications in the real situation are being ignored. However, the syntactic aspects of the Dutch situation that are suppressed here so far do not lend themselves to supporting any kind of stringset argument. In some cases, they make a valid stringset argument even harder to achieve. For example, Zaenen (1979) points out that a verb like *wil* does not always occur before the nonfinite verbs with which it co-occurs. Beside the order in (34a), for instance, the order in (34b) is also grammatical.

(34) (a) dat   Jan met Marie in de tuin   wil spelen
          that Jan with Marie in the garden will play

(b) dat   Jan met Marie in de tuin   spelen wil
     that Jan with Marie in the garden play   will
     'that Jan will play with Marie in the garden'

This additional variation makes it much harder to begin setting up an argument involving a formal cross-serial dependency structure of the type $a_1a_2 \ldots a_nb_1b_2 \ldots b_n$, which is what Huybregts (1976) purports to do (but does not do). The task of setting up such an argument is similarly made harder by the fact that many transitive verbs can be used intransitively, so that it will not necessarily cause ungrammaticality if an *NP* is removed from a well-formed example. The reader can verify this by experiment. It is in fact quite difficult to construct examples that convincingly change their grammaticality when random *NP*'s or verbs are inserted or removed, since so many strings turn out to be grammatical by accident under an unintended interpretation, or ungrammatical by virtue of extraneous factors that do not relate to the alleged dependency between *NP*'s and verbs.

It is difficult to tell at this stage whether some CFL-inducing theory of grammar might be capable of providing for an adequate grammar for Dutch and a semantic interpretation system to go with it. However, that is not the claim at issue. We are concerned here with the thesis that Dutch, as a stringset, is not a CFL. We have said enough to show that there is no reason at all to believe that claim.

6. MOHAWK

The best known and most influential argument for the non-context-freeness of a natural language is due to Postal (1964). Postal argues that the interaction of the processes of nominalization and incorporation in Mohawk (a Northern Iroquoian language of Quebec and upper New York state) yields a property that places Mohawk outside the CFL's. In brief, a verb in Mohawk can incorporate a noun-stem, i.e. it can have the internal structure [v Prefixes Noun-stem Verb-stem Suffixes]. It incorporates the noun-stem of its subject if it is intransitive, or the noun-stem of its direct object if it is transitive. Thus we have sentences consisting of single words with morpheme glosses like 'it-tree-stands' or 'he-meat-eats' and meanings like 'The tree stands', 'He eats meat'. But a verb like 'meat-eat' can be nominalized to make a noun meaning 'meat-eating', and this noun could itself be incorporated into a verb, which could be nominalized, and so on. Hence Mohawk has an infinite set of noun-stems. But sometimes a verb with an incorporated subject (object) noun-stem

occurs with an overt subject (object) *NP*. In such cases the noun-stem in the verb must exactly match the noun-stem in the external *NP*. This is string-copying over an infinite set of strings (the set of noun-stems), hence Mohawk is an *xx* language and not a CFL.

Postal's paper has been criticized in the subsequent literature, but the critiques have had little impact. The best known critique is Reich (1969). Of this paper, Levelt (1974, 21n) says: "The large number of essential errors in this article ... gives rise to some doubt as to the carefulness of the editors of *Language*". Fittingly, Reich's paper (which includes, among other things, an assertion without proof that the FSL's are not a subset of the CFL's) has been generally ignored by linguists.

There are certainly mathematical flaws in Postal's argument the way he originally gave it. The major one is that Postal relies on the following as his chief mathematical premise:

(35)   It has been proven by Chomsky that the language consisting of all and only the strings [*XX*] is not a context-free language, where *X* varies over an infinite set of strings in an alphabet of two or more symbols. (1964, p. 146)

The theorem Postal alludes to was never proven by Chomsky, and could not be, because it is not a theorem at all. There are languages of the type specified that are CFL's, as noted by Daly (1974, p. 68) and Elster (1978, p. 47).[14] An additional flaw is that Postal assumes:

(36)   It can be demonstrated that Mohawk lies outside the bounds of context-free description by showing that it contains, as a subpart, an infinite set of sentences with formal properties of the language [*XX*]. (1964, p. 147)

This is not correct, as can be readily seen by noting that the non-CFL $\{xx \mid x \in L((a + b)^+)$ is an infinite proper subpart of the context-free (and finite state) language denoted by the expression $(a + b)^+$. This formal inadequacy is commented on by Daly (1974) and also by Fidelholtz (1975, p. 496, fn. 2).

However, these mathematical slips are unimportant, because they have been remedied by Langendoen (1977, who characterized precisely the conditions under which languages of the sort alluded to in (35) are non-context-free, and shows how to give a formally valid proof that Mohawk is non-context-free by intersecting it with a FSL to obtain a provably non-context-free *xx* language. He does this as follows. First he

sets up the following abbreviations:

(37) a = the translation into Mohawk of 'the man'
         b = the translation into Mohawk of 'admired'
         c = the translation into Mohawk of 'liking (of)'
         d = the translation into Mohawk of 'praising (of)'
         e = the translation into Mohawk of 'house'

Then he defines the finite state language $F = a(c + d)^* eb(c + d)^* e$. F contains all possible strings having the form *the man x-house-admired y-house* where x and y are arbitrary strings with meanings such as 'praising-of-liking-of-liking-of-praising-of-praising-of-liking-of' and so on. Postal's claim about Mohawk is that sentences from F will only be grammatical Mohawk sentences if the first string from $(c + d)^*$ is identical to the second (i.e. if x is the same as y), so Langendoen intersects F with Mohawk and obtains the language $L^m = \{axebxe \mid x \in L((c + d)^*)\}$. This is an xx language and thus not a CFL. Hence Mohawk is not a CFL.

Langendoen's version of the argument ignores one point that Postal mentions: there is a requirement "that a Modifier constituent precede the noun whose stem is to be doubled" (1964, p. 147, fn. 29). This is not a trivial point, because if the modifier (demonstrative, possessive, or whatever) is essential, as Postal claims, then F does not intersect with the set of stem-doubling sentences in Mohawk at all. Assume, then, that we take account of this point by adding a Mohawk modifier such as *thik*∧ 'that' after b in the specification of F. There are then no flaws in the argument at all. We only need the crucial empirical premises about what is grammatical in Mohawk (and thus what is in L^m) to be sound.

Reich (1969, p. 833) correctly summarizes the three empirical premises as follows:

The first is that the acceptable [read 'grammatical' – GKP/GJMG] sentences of Mohawk include ... sentences where the incorporated object co-occurs with an identical direct object which has been preceded by a modifier. The second ... is that there are no Mohawk sentences of this type in which the object of the verb and the incorporated object are not identical; that is, there are no Mohawk sentences of this type which are in the form $XY[X \neq Y$ – GKP/GJMG]. The third ... is that the nouns which are used as objects (both incorporated and free-standing) can be arbitrarily long, because of a process of making a noun out of a verb containing an incorporated noun.

The first premise should not be in doubt. Postal discusses the object-doubling construction in his dissertation (Postal (1962); see pp. 290ff) and gives elicited examples like (38).[15]

(38)   i?i k**nuhs**nuhwe?s *thik*∧ ka**nuhs**a
         I   *house*-like   that   *house*
         'I like that house.'

Those anxious to be assured that this construction occurs in spontaneous text need look no further than Michelson (1980, 29, line 21), where the stem -ri(h)w- 'matter' occurs incorporated in a verb with the stem -yʌ- 'put' and also in its object NP.

Reich's main attack is based on rebutting the third premise. He cites Floyd Lounsbury as claiming that "you never get more than one incorporation at the creative (syntactic) level", i.e. if an incorporated noun-stem itself contains an incorporated noun-stem, the latter is a frozen part of "a common idiom, a lexeme in the language" (1969, p. 834). Thus, Lounsbury is claiming, there is no productive feeding of incorporation by nominalization and conversely; the closest approach to this comes when incorporation happens to involve a noun stem which contains the lexically frozen remains of an incorporation which is not synchronically the result of the syntactic incorporation rule. This may be so,[16] but we are not inclined to take up this line of argument. As already noted (see Note 13), we do not necessarily accept that the lexicon of a natural language has to be a finite list, so a version of Langendoen's argument might be constructible even if Lounsbury were right. Moreover, we certainly do not follow Reich in his view that if the longest incorporated noun-stem ever observed in a text had, say, only three layers of repeated incorporation, the correct grammar should set a three-layer limit. We accept the standard idealization, with its corollary that the set of attested and likely-to-be-attested sentences does not exhaust the set of grammatical sentences.

We believe that the Achilles' heel of the Mohawk argument lies in the second premise mentioned in the quote above; but no one has previously shown this. Reich mentions that Lounsbury has 'reservations' about the second premise, and gives translations of, but does not exhibit, three sentences which Lounsbury claims violate the identity requirement. He does not go into any detail. Postal acknowledges in his paper that there are other incorporation constructions than the one he is talking about:

There are certain other cases where a verb contains an incorporated noun stem which does not match the following external noun stem. These are due to minor rules and do not affect the present discussion. (1964, p. 148, fn. 30)

What has to be shown if Postal's argument is to be impugned is that the other cases referred to *do* affect the present discussion, and affect the soundness of Langendoen's reformulation of the argument.

One example of the sort of construction Postal was referring to is the construction generally known as *classificatory incorporation* in the literature on Iroquoian (see, e.g., Woodbury, 1975). These are mentioned in footnote 19 on pp. 405–406 of Postal (1962). Postal's analysis of

incorporation makes reference to a morpheme called 'inc' which is replaced by a noun-stem when the incorporation transformation ('T-incorporation') applies, or is spelled out in one of various forms as an empty morph (cf. Lounsbury, 1953, p. 75) if incorporation does not apply. In the footnote cited he states that in addition to what T-incorporation does:

> there appear to be a few special rules which replace the inc marker by certain noun stems before T-incorporation applies. One such rule inserts the stem *naskw* 'animal' when the object noun of the sentence is a member of Noun Stem_mammal. Thus we would get such sentences as
>
>   sawatis hranaskwhinu? ne yaoahkwari 'John will buy a bear'
>
> And there are other cases where a stem with a more general meaning is inserted. I have found cases where the stem for 'berry' *hy* is used with fruits, where the stem for 'water' *hnek* is used with drinkable liquids, etc.

Marianne Mithun (personal communication) supplies a typical example of this sort:

(39)   wa?k*hnek*ahnínu? ne otsì?tsa?
        I-*liquid*-bought   flower/wine
        'I bought the wine.'

Here the meaning of the incorporated stem resolves the ambiguity of the external stem -*tsi?ts*-, which means both 'flower' and 'wine'. Other examples are supplied by Bonvillain (1974, pp. 21–22); e.g. (40).

(40)   wake*selehta*hni:nu?se? ne? 'bike'
        I her-*vehicle*-bought   bike
        'I bought her a bike.'

These and similar cases certainly show that there are instances of incorporated noun-stems in transitive verbs that fail to match the noun-stems of their direct object NP's so that Postal's argument as originally set out is invalidated empirically as well as formally. But when we examine the form of the Langendoen version of Postal's case (hence-forth the LP argument), we find that provided classificatory incorporation is limited to a finite list of simple stems (as Postal strongly implies it is), the construction has no implications for the LP argument at all. It is far too simplistic to assume that *any* discovery of stem-matching failure will entail irreparable failure of the whole argument.

   Consider another kind of case, alluded to in Reich's discussion: null head nouns in subject or object NP's with modifiers. Postal (1962, p. 395) cites the following pair of sentences as synonymous:

(41) (a) kan*uhs*rak∧   think∧ kan*uhs*a
         it-*house*-white   that *house*
         'That house is white.'
     (b) kan*uhs*rak∧   think∧
         it-*house*-white   that
         'That house is white.'

Here an intransitive subject is seen with and without its head noun. The same possibility is present for a direct object, as shown by this example from Bonvillain and Francis (1980, p. 80, line 27; morpheme segmentation on p. 88):

(42)   keli?wa*nú*:we?s ki?
        I it-*idea*-like   this
        'I agree to this.'

Again we appear to have a counterexample to the claim that incorporated stems match external stems. But if the LP argument as summarized above is examined closely, it will be found that the argument is only concerned with those sentences that are in Mohawk and have glosses of the form '(the) man x-house-admired that y-house' (correcting for the presence of the modifier as mentioned above), where x and y are noun-stems formed by using incorporation and nominalization to combine 'praise' and 'like. Strings that do not end with the form meaning 'house' are not relevant at all. Thus the LP argument is fully robust enough to stand up to Reich's criticism, insofar as the evidence he offers is concerned.

However, there is a construction in Mohawk that has much more serious implications for the LP argument. Essentially all the relevant facts are to be found in Postal (1962), though we shall cite one or two other sources as well. We shall call the construction at issue the *possessed incorporation* construction. Briefly, when the subject or direct object *NP* of a verb contains a possessive *NP* modifier, it is possible to incorporate the noun-stem denoting the possessed entity, keeping the external *NP*; but it is also possible to drop the possessed noun from the external NP so that the possessor noun constitutes the whole of the subject or direct object NP. In the latter case the verb agrees with the possessor. These facts are illustrated in (43).

(43)(a) i?i k-nuhwe?s ne ka-*nuhs*-a?
        I   like   *house*
        'I like the house.' (Postal 1962, p. 283, E147; morpheme gloss added.)

(b)   i?i k-nuhwe?s ne sawatis hra-o-*nuhs*-a?
      I   like   John('s)   *house*
      'I like John's house.' (Postal 1962, p. 321, E243; morpheme gloss
      added.)

(c)   i?i k-*nuhs*-nuhwe?s ne sawatis hrao-*nuhs*-a?
      I   *house- like*   John('s)   *house*
      'I like John's house.' (Postal, 1962, p. 291 E168; morpheme
      gloss added.)

(d)   i?i hrai-*nuhs*-nuhwe?s ne sawatis
      I   *house*-like   John
      'I like John's house.' (Postal, 1962, p. 320, E239; morpheme
      gloss added.)

(e)   *i?i hrai-nuhwe?s ne sawatis hrao-*nuhs*-a?
      I   like   John('s)   *house*
      (Postal, 1962, p. 321, E244; morpheme gloss added.)

In (40a) we see a simple example with subject, verb, and object. (The
subject pronoun *i?i* 'I' would normally be dropped.) A similar sentence
with a possessive *NP* in the object phrase is shown in (40b). In (40c) the
noun-stem denoting the possessed item, i.e. the head of the object *NP*,
has been incorporated, but is also present in the object *NP*. In (40d) the
head noun has been dropped from the object *NP*, leaving the in-
corporated noun-stem as the only realization of the notion 'house', and
now the verb shows a different agreement pattern: as (40e) illustrates,
the prefix *hrai*- is not correct for sentences with object *NP*'s like 'John's
house'. It is the appropriate prefix for sentences like 'I like John' (see,
e.g., Postal's sentence E159 (1962, p. 285) for a similar case), and this is
the prefix taken by the verb in (40d).

An example like (40d) clearly shows an incorporated noun-stem that
fails to match the noun-stem of the direct object of its host verb, which
crucially must not occur if Postal's overall argument is to be sound.[17] It
seems clear that in constructing the argument he assumed that additional
considerations of detail like the agreement pattern seen in (40d) would
suffice to isolate the possessed incorporation construction (his 'posses-
sor agreement' construction: 1962, p. 319) so that it could be set aside as
irrelevant. But he has not provided a demonstration that this can be
done. It seems to us that the LP argument does not survive once
possessed incorporation sentences are brought into consideration. What
is crucial for the LP argument is that sentences of Mohawk with glosses
like 'The man praising-of-liking-of-house-admired that liking-of-praising-
of-house' should be ungrammatical. But it is clear from the above

discussion that such sentences will not be ungrammatical at all; they will merely have absurd meanings as their only possible readings – in the case just cited, the meaning will be 'The man admired that liking-of-praising-of-house's praising-of-liking-of-house'. Verb agreement facts do not affect this conclusion in any way, for the abstract head nouns meaning 'liking-of-praising-of-house' and 'praising-of-liking-of-house' will determine the same agreement prefixes. The contrast between (40c) and (40d) only arises because of the contrast in agreement class memberships between the masculine human noun *sawatis* 'John' and the nonhuman noun *-nuhs-* 'house'. In the type of sentence that figures in the LP argument, the noun stems we are concerned with are always, and crucially, abstract nouns built up by iteration of the operations of incorporation and nominalization.

Thus the construction we call possessed incorporation demonstrates that the intersection of Langendoen's language F with Mohawk will in fact not be of the form $\ldots x \ldots x \ldots$ where $x \in L(c + d)^*)$, but rather of the form $\ldots x \ldots y \ldots$, where x and y are drawn from $L((c + d)^*)$ but are not necessarily identical. This is just another way of saying that Mohawk has F, a FSL, as an infinite subset. Nothing whatever follows about the context-freeness of Mohawk, and thus the LP argument against Mohawk being a CFL fails like all the previous arguments.

7. CONCLUSIONS

Notice that this paper has not claimed that all natural languages are CFL's. What it has shown is that every published argument purporting to demonstrate the non-context-freeness of some natural language is invalid, either formally or empirically or both.[18] Whether non-context-free characteristics can be found in the stringset of some natural language remains an open question, just as it was a quarter century ago.

Whether the question is ultimately answered in the negative or the affirmative, there will be interesting further questions to ask. If it turns out that natural languages are indeed always CFL's, it will be reasonable to ask whether this helps to explain why speakers apparently recognize so quickly whether a presented utterance corresponds to a grammatical sentence or not, and to associate structural and semantic details with it. It might also be reasonable to speculate about the explanation for the universally context-free character of the languages used by humans, and to wonder whether evolutionary biological factors are implicated in some way (Sampson (1979) could be read in this light). And naturally, it will be reasonable to pursue the program put forward by Gazdar (1981,

in press) to see to what extent CFL-inducing grammatical devices can be exploited to yield insightful descriptions of natural languages that capture generalizations in revealing ways.

If a human language that is not a CFL is proved to exist, on the other hand, a different question will be raised: given the non-context-free character of human languages in general, why has this property been so hard to demonstrate that it has taken over twenty-five years to bring it to light since the issue was first explicitly posed? If human languages do not have to be CFL's, why do so many (most?) of them come so close to having the property of context-freeness? And, since the CFL's certainly constitute a very broad class of mathematically natural and computationally tractable languages, what property of human beings or their communicative or cognitive needs is it that has caused some linguistic communities to reach beyond the boundaries of this class in the course of evolving a linguistic system?

Either way, we shall be interested to see our initial question resolved, and further questions raised. One cautionary word should be said, however, about the implications (or lack of them) that the answer will have for grammatical studies. Chomsky has repeatedly stated that he does not see weak generative capacity as a theme of central importance in the theory of grammar, and we agree. It is very far from being the case that the recent resurgence of interest in exploring the potential of CF-PSG or equivalent systems will, or should, be halted dead in its tracks by the discovery (if it is ever forthcoming) that some natural language is not a CFL. In the area of parsing, for instance, it seems possible that natural languages are not only parsed on the basis of constituent structure such as a CF-PSG would assign, but are parsed as if they were finite state languages (see Langendoen (1975) and Church (1980) for discussion along these lines). That is, precisely those construction-types that figure in the various proofs that English is not an FSL appear to cause massive difficulty in the human processing system; the sentences crucial to the proofs are for the most part unprocessable unless they are extremely short (yet the arguments for English not being an FSL only go through if length is not an issue). This means that in practice properties of finite state grammars are still of great potential importance to linguistic theory despite the fact that they do not provide the framework for defining the total class of grammatical sentences. The same would almost certainly be true of CF-PSG's if they were shown to be inadequate in a similar sense. It is highly unlikely that the advances made so far in far in phrase structure description could be nullified by a discovery about weak generative capacity. Moreover, there are known

to be numerous ways in which the power of CF-PSG's can be marginally enhanced to permit, for example, xx languages to be generated without allowing anything like the full class of recursively enumerable or even context-sensitive languages (see Hopcroft and Ullmann (1979, Chapter 14) for an introduction to this topic, noting especially Figure 14.7 on p. 393. The obvious thing to do if natural languages were ever shown not to be CFL's in the general case would be to start exploring such minimal enhancements of expressive power to determine exactly what natural languages call for in this regard and how it could be effectively but parsimoniously provided in a way that closely modelled human linguistic capacities.

In the meantime, it seems reasonable to assume that the natural languages are a proper subset of the infinite-cardinality CFL's, until such time as they are validly shown not to be.

University of California
University of Sussex

NOTES

[1] Note that although a language is by definition a CFL if there is a context-free phrase structure grammar (CF-PSG) that generates it, it does not follow that if a grammar G generates a CFL, G is a CF-PSG (though it does follow that a weakly equivalent CF-PSG must exist). A grammar of any arbitrary type might happen to generate a CFL. And for all we know, it might be that arc pair grammars or 'government-binding' grammars *always* do.

[2] Another is that if PS rules are unsuccessful on their own, then a new type of rule must be introduced. See Gazdar *et al.* (1981) for a CF-PSG that captures the generalization Akmajian and Heny are concerned with by means of a device for collapsing rules rather than a new rule type.

³ See Gazdar (in press) for a brief discussion of how subject-verb agreement facts can be elegantly described by collapsing sets of CF-PSG rules.

⁴ From this point on, any unexplained notation for representing grammars, expressions, or languages will be taken from Hopcroft and Ullmann (1979), see especially pp. 28–29 and 79–80. Note in particular that we write regular expressions (for representing FSL's) in boldface; A^* means the set of all strings of members of A; A^+ means all the nonempty members of A^*; $L(\varphi)$ means the language denoted by φ; and we use the notation '$A \to B \mid C$' to mean 'A rewrites as either B or C'. We also use the standard abbreviation '$A \to B(C)D$' to mean 'A rewrites as either BCD or BD'.

⁵ An equivalent demonstration is to show that the fragment of English involved is representable by a regular expression:

> ((**Which problem did your professor say ((she + you) thought)* was unsolvable) + (Which problems did your professor say ((she + you) thought)* were unsolvable))**

⁶ Selkirk does not specifically say that she means inadequacy as regards weak generative capacity, of course. She might mean that it has been demonstrated that CF-PSG's cannot assign appropriate structural descriptions to sentences. But in fact this claim has not been demonstrated either, and it is much harder to investigate rigorously. While matters of weak generative capacity (ability to generate stringsets) are relatively well understood in mathematical terms, notions like 'strong generative capacity' (never formally explicated in the literature, but having something to do with generating the 'right' tree-set for a given stringset) are far less amenable to formal work, because issues about what trees to assign to what strings depend on much more subtle and indirect arguments and evidence. If generally current transformationalist views on what surface structures are like are assumed, then for a well-studied language like English, CF-PSG does rather well (see e.g. Gazdar (1981)).

⁷ What would count as identity here is rather hard to define precisely, because the identity relation must ignore the presence of the inflectional comparative suffix *-er*, and also has to ignore morphologically determined suppletion, as seen in (i), which must, presumably, be counted ungrammatical in the same way as (1) despite its superficially nonidentical adjective stems.

(i)   This one is better than that one is GOOD.

⁸ We would not have succeeded in constructing this grammar without the help of Aravind Joshi, Stanley Peters, and Robert Ritchie, who told us it was possible and gave us hints on how to start looking for the grammar in (9).

  Note that the complement of (8) in $\alpha(a + b)^*\beta(a + b)^*\gamma$ is not a CFL but an xx language. However, the CFL's are not closed under complementation (Hopcroft and Ullmann, 1979, pp. 134–135), so this is not a perplexing fact.

⁹ English comparative clauses can in fact be described fairly elegantly with a CF-PSG; see Gazdar (1980).

¹⁰ This assumption is, of course, crucial to Corstius's $a^n b^n c^n$ version of the *respectively* argument (see Levelt (1974, pp. 31–32)). If, as we maintain, the assumption is false, then his argument does not go through.

¹¹ McCawley (1968, pp. 164, 168) cites examples which illustrate this point.

¹² Wachtel (1981) argues that language-particular gender/number assignments give rise to a language-imposed classification of referents in cases of pragmatically controlled anaphora. His position, taken together with de Cornulier's metalinguistic analysis of sentences like (22a) and (22b), would lead us to expect sentences of the form shown in (i) and (ii) in a language with gender indications *-a*, *-b*, and *-c*.

(i)   In my view, oranges-*a*, lemons-*b*, and bananas-*c* are, respectively, delicious-*a*, bitter-*b*, and fattening -*c*.

(ii)   In my view, oranges-*a*, lemons-*b*, and bananas-*c* are, in the order in which I cite them, delicious-*a*, bitter-*b*, and fattening-*c*.

So, even in such a language, it would be neither necessary nor desirable to impose a purely syntactic matching condition on sentences like (i).

[13] One line we think worth exploring would be based on the notion of the lexicon as an infinite set of forms generated by a recursive procedure. It is not really in doubt that something along these lines will in due course have to be developed. As explicitly noted by Langendoen (1981), there are infinite sets of related words in English – for example, number names, and recursively constructed ancestor terms like *great-great-grandfather*. (Interestingly, Langendoen observes that no non-context-free or even non-finite-state sets of words can be found in currently known languages, though such sets can readily be invented. The lexicon of a human language may be infinite, but will apparently always be an FSL.) What needs to be examined is whether Dutch could be said to have an indefinitely extensible set of verbs with meanings like 'see write', 'let see write', 'help let see write', etc. It is a moot point whether a system could be set up to provide for an infinite set of internally complex members of the category V, and associate the members syntactically and semantically with the appropriate number of NP arguments, and still be CFL-inducing. (If it were not CFL-inducing, it would be an excessively powerful system along the relevant parameter, because it seems clear that all we are trying to do is to assign appropriate structural descriptions to a context-free set of sentences.)

[14] Daly's example is $\{ab^nab^n \mid n \geq 0\}$. A CF-PSG with the rules $S \to aZ$, $Z \to bZb \mid a$ generates it.

[15] The following points are relevant to any careful study of the Mohawk examples. We shall be quoting Mohawk examples both in Postal's fairly abstract phonological representation, as in (38), and in the phonemic notations used by other Iroquoianists. Agreement on spelling of morphemes should not therefore be expected (though the spelling of the mid-central nasal vowel has been silently normalized to ʌ throughout). Mohawk has a rich morphophonemics, so stems and other formatives will appear in varying shapes even within one example in one transcription system. Incorporated noun-stems and their glosses will be italicized for the reader's convenience, but morpheme segmentation and glossing is in general kept to a minimum. Unexplained prefixes, infixes, and suffixes are usually agreement inflections, epenthetic vowels, and markers of aspect respectively. The particle *ne* which appears in later examples and is ignored in glossing is commonly translated as 'the' in the literature. It occurs with proper names as well as common nouns. The proper name *sawatis* is always glossed as 'John'; it is the Mohawk transliteration of the French *Jean-Baptiste* 'John the Baptist'.

[16] There are a mass of lexical restrictions on incorporation: verbs stems that do not allow any incorporation, noun-stems that cannot be incorporated, verb-stems that are required to have an incorporated noun-stem, and perhaps some noun-stems that only occur incorporated (see Postal, 1962, p. 286). Postal mentions that the verbs that permit no incorporation may be in a majority. Such lexical idiosyncrasy tends to increase the plausibility of the claim that Reich attributes to Lounsbury.

[17] It is worth remarking that examples like (40d) are not marginal or peculiar in Mohawk. With an inalienably possessed item as the incorporated element they are extremely common; in Bonvillain and Francis (1980) there are a number of examples like these:

kwískwis ya?tho?n*yukwa*íhshta?
pig   he it-*snout*-grabbed
'he grabbed the pig's nose' (line 60, p. 84; morpheme segmentation on p. 92)
saho*hna?ts*li?khú:ne? ne kwískwis
while-he it-*ass*-bite the pig

sahohna?tsli?khú:ne? ne kwískwis
while-he it-ass-bite the pig
'while he kept biting the pig's ass' (line 68, p. 85; morpheme segmentation on
p. 93).

But alienably possessed stems occur in this construction too, as seen in (43d) in the text. Postal confirms this point, and has checked it with another Mohawk specialist (Karin Michelson).

[18] After this paper was completed, Arnold Zwicky reminded us of the existence of his (1963) paper on this topic, which we had overlooked. The most convincing example discussed there involves the names for very large numbers in English. Assume that the largest number that has a one-word name is named *zillion*. Then the square of this number has to be called (*one*) *zillion zillion*. An even larger number is *one zillion zillion, one zillion, and one*. But **one zillion, one zillion zillion, and one* is not a legal name for this number, or for any number. In general, the powers of zillion must be given in such a way that $zillion^n$ follows $zillion^{n+1}$ for all n. But a language of the general form

$$\{pz^{n_1}pz^{n_2}\ldots p^{n_i} \mid (\forall_j, 1 \leqslant j \leqslant i)(n_j > n_{j+1})\}$$

is not context-free, as Zwicky shows. It follows that there is an infinite set of number names in English that is not context-free (and it is extractable by means of intersection with a regular set). The interest of this argument in the context of the study of natural languages is, however, greatly lessened by the fact that it deals with the internal structure of elements of a representational system for mathematics. We would maintain that knowledge of how to construct such number names (which, of course, has to be explicitly taught to children who speak English perfectly well) is knowledge of mathematics rather than of language.

REFERENCES

Aho, A. V. and J. D. Ullmann: 1972, *The Theory of Parsing, Translation and Compiling*, Volume I: *Parsing* (Prentice-Hall, Englewood Cliffs, New Jersey).

Aho, A. V. and J. D. Ullmann: 1973, *The Theory of Parsing, Translation and Compiling*, Volume II: *Translation and Compiling* (Prentice-Hall, Englewood Cliffs, New Jersey).

Akmajian, A. and F. Heny: 1975, *An Introduction to the Principles of Transformational Syntax* (MIT Press, Cambridge, Mass.).

Allerton, D. J.: 1980, *Essentials of Grammatical Theory* (Routledge and Kegan Paul, London).

Bach, E.: 1974, *Syntactic Theory* (Holt Rinehart and Winston, New York).

Bar-Hillel, Y. and E. Shamir: 1960, 'Finite State Languages: Formal Representations and Adequacy Problems', reprinted in Y. Bar-Hillel (1964), *Language and Information* (Addison-Wesley, Reading, Mass.), pp. 87–98.

Bonvillain, N.: 1974, 'Noun Incorporation in Mohawk', in M. K. Foster (ed.), *Papers from the 1972 Conference on Iroquoian Research* (National Museum of Man, Ottawa, Canada), pp. 18–26.

Bonvillain, N. and B. Francis: 1980, 'The Bear and the Fox, in Akwesasne Mohawk', in Mithun and Woodbury (eds.), pp. 77–95.

Bresnan, J. W.: 1976, 'Evidence for a Theory of Unbounded Transformations', *Linguistic Analysis* 2, 353–393.

Bresnan, J. W.: 1978, 'A Realistic Transformational Grammar', in M. Halle, J. W. Bresnan, and G. A. Miller (eds.), *Linguistic Theory and Psychological Reality* (MIT Press, Cambridge, Mass.).

Chomsky, N.: 1956, 'Three Models for the Description of Language', *I.R.E. Transactions on Information Theory*, Volume IT-2, Proceedings of the Symposium on Information

Theory, September, pp. 113–123.

Chomsky, N.: 1963, 'Formal Properties of Grammars', in R. D. Luce, R. R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, Volume II (John Wiley, New York).

Chomsky, N.: 1977, 'On *Wh*-Movement', in P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax* (MIT Press, Cambridge, Mass.).

Chomsky, N.: 1981, *Lectures on Government and Binding* (Foris, Dordrecht).

Church, K. W.: 1980, *On Memory Limitations in Natural Language Processing*, MSc Dissertation, MIT, Cambridge, Mass.

Cornulier, B. de: 1973, 'But If "Respectively" Meant Something?', *Papers in Linguistics 6*, 131–134.

Culicover, P. W.: 1976, *Syntax* (Academic Press, New York).

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

Elster, J.: 1978, *Logic and Society: Contradictions and Possible Worlds* (John Wiley, New York).

Fidelholtz, J.: 1975, Review of J. Kimball, *The Formal Theory of Grammar, Language 51*, 493–499.

Fodor, J. A.: 1975, *The Language of Thought* (Thomas Crowell, New York).

Fromkin, V. and R. Rodman: 1978, *An Introduction to Language*, Second Edition (Holt Rinehart and Winston, New York).

Gazdar, G.: 1980, 'A Phrase Structure Syntax for Comparative Clauses', in T. Hoekstra, H. van der Hulst and M. Moortgat (eds), *Lexical Grammar* (Foris Publications, Dordrecht), pp. 165–179.

Gazdar, G.: 1981, 'Unbounded Dependencies and Coordinate Structure', *Linguistic Inquiry 12*, 155–184.

Gazdar, G.: (in press). 'Phrase Structure Grammar', in P. Jacobson and G. K. Pullum (eds.), *The Nature of Syntactic Representation* (D. Reidel, Dordrecht, Holland).

Gazdar, G., G. K. Pullum, and I. A. Sag: 1981, *Auxiliaries and Related Phenomena in a Restrictive Theory of Grammar* (Indiana University Linguistics Club, Bloomington, Ind.).

Green, G. M.: 1971, 'Unspeakable Sentences, Book I', *Linguistic Inquiry 2*, 560.

Grinder, J. T. and S. H. Elgin: 1973, *Guide to Transformational Grammar* (Holt Rinehart and Winston, New York).

Hinton, G.: 1978, 'Respectively Reconsidered', *Pragmatics Microfiche 3.3*, 912–914.

Hopcroft, J. and J. D. Ullmann: 1979, *Introduction to Automata Theory, Languages and Computation* (Addison-Wesley, Reading, Mass.).

Hurford, J. R.: 1980, 'Generative Growing Pains', *Lingua 50*, 117–153.

Huybregts, M. A. C.: 1976, 'Overlapping Dependencies in Dutch', *Utrecht Working Papers in Linguistics 1*, 24–65.

Johnson, D. E. and P. M. Postal: 1980, *Arc Pair Grammar* (Princeton University Press, Princeton, N.J.).

Kimball, J.: 1973, *The Formal Theory of Grammar* (Prentice-Hall, Englewood Cliffs, New Jersey).

Langendoen, D. T.: 1975, 'Finite State Parsing of Phrase-Structure Languages and the Status of Readjustment Rules in the Grammar', *Linguistic Inquiry 6*, 533–554.

Langendoen, D. T.: 1977, 'On the Inadequacy of Type-3 and Type-2 Grammars for Human Languages', in P. J. Hopper (ed.), *Studies in Descriptive and Historical Linguistics: Festschrift for Winfred P. Lehmann* (John Benjamin, Amsterdam, Holland), pp. 159–171.

Langendoen, D. T.: 1981, 'The Generative Capacity of Word-formation Components', *Linguistic Inquiry 12*, 320–322.

Levelt, W. J. M.: 1974, *Formal Grammars in Linguistics and Psycholinguistics*, Volume II: *Applications in Linguistic Theory* (Mouton, The Hague, Holland).

Lounsbury, F.: 1953, *Oneida Verb Morphology* (Yale University Publications in Anthropology, No. 48, Yale University, New Haven, Conn.).

McCawley, J. D.: 1968, 'The Role of Semantics in a Grammar', in E. Bach and R. T. Harms (eds.), *Universals in Linguistic Theory* (Holt Rinehart and Winston, New York).

Michelson, K.: 1980, 'Mohawk Text: The Edge of the Forest revisited', in Mithun and Woodbury (eds.), pp. 26–40.

Mithun, M. and H. Woodbury (eds.): 1980, *Northern Iroquoian Texts* (IJAL Native American Texts Series, No. 4) (University of Chicago Press, Chicago/University Microfilms International, Ann Arbor, Mich.).

Pinker, S.: 1979, 'Formal Models of Language Learning', *Cognition* 7, 217–283.

Postal, P. M.: 1962, *Some Syntactic Rules in Mohawk*, Doctoral Dissertation (Yale University, New Haven, Conn.; Published by Garland, New York, 1979).

Postal, P. M.: 1964, 'Limitations of Phrase Structure Grammars', in J. A. Fodor and J. J. Katz (eds.), *The Structure of Language: Readings in the Philosophy of Language* (Prentice-Hall, Englewood Cliffs, N.J.), pp. 137–151.

Reich, P. A.: 1969, 'The Finiteness of Natural Language', *Language* 45, 831–843.

Sampson, G.: 1975, *The Form of Language* (Weidenfeld and Nicolson, London).

Sampson, G.: 1979, 'A Non-Nativist Account of Language Universals', *Linguistics and Philosophy* 3, 99–104.

Selkirk, E. O.: 1977, 'Some Remarks on Noun Phrase Structure', in P. W. Culicover, T. Wasow and A. Akmajian (eds.), *Formal Syntax* (Academic Press, New York), pp. 285–316.

Wachtel, T.: 1981, 'Sex and the Single Pronoun', unpublished paper (University of Warsaw).

Winograd, T.: 1972, *Understanding Natural Language* (Academic Press, New York). Also *Cognitive Psychology* 3, No. 1 (1972).

Woodbury, H.: 1975, 'Onondaga Noun Incorporation: Notes on the Interdependence of Syntax and Semantics', *International Journal of American Linguistics* 41, 10–20.

Zaenen, A.: 1979, 'Infinitival Complements in Dutch', *Papers from the Fifteenth Regional Meeting* (Chicago Linguistic Society, Chicago, Ill.), pp. 378–389.

Zwicky, A. M.: 1963, 'Some Languages That Are Not Context-free', *Quarterly Progress Report of the Research Laboratory of Electronics, MIT*, 70, 290–293.