CANONICAL LANGUAGES

with special study of
a canonical predicate calculus for number theory
a canonical grammar for the English language

by

Donald Kingsbury BSc

A thesis submitted to the Faculty of Graduate Studies and research in partial fulfilment of the requirements for the degree of Master of Science.

Department of Mathematics
McGill University
Montreal

August 1959

INTRODUCTION

A language is in canonical form if its basis has the following structure due to Post:

(i) There exists a finite alphabet of distinct letters which can be juxtaposed to form strings, ie, any letter is a string, if b and c are strings then bc is a string, nothing else is a string.

(ii) There exists a special finite set of strings called primitive assertions.

(iii) There exists a finite set of productions each of the following form:

$$\begin{array}{l}\left.\begin{array}{llllll} g_{11} & P_{11} & g_{12} & P_{12} & \cdots & g_{1n} \\ g_{21} & P_{21} & g_{22} & P_{22} & \cdots & g_{2n'} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ g_{K1} & P_{K1} & g_{K2} & P_{K2} & \cdots & g_{Kn^{(K)}} \end{array}\right\}\text{data} \\ \hline \left.\begin{array}{llllll} g_{1} & P_{J_1} & g_{2} & P_{J_2} & \cdots & g_{K} \end{array}\right\}\text{product} \end{array}$$

where the g's are given strings, possibly null, and the P's represent operational variables of the production chosen from a finite set of such string variables, possibly including the null string. We make the restriction that each P_j from the product must be chosen from the P_{1m} of the data, that each datum and the product must contain at least one P, and that no choice of P's can be allowed which would permit a null product.

The strings consisting of the primitive assertions, and all the strings obtainable by repeated use of the productions starting with the primitive assertions, will be called the assertions of the system. Any datum of such a production will be the product of another production, or a primitive assertion.

Post originally developed this definition of language as an abstract structure from which he hoped to prove theorems about languages in general, especially concerning their decision problems. In his 1941 paper (1) he proved his remarkable normal form theorem which states that every canonical language can be reduced to a normal language, ie one which has but a single primitive assertion with each of its productions of the form $gP \longrightarrow Pg'$. Thus any theorem concerning the structurally simpler normal form is also a theorem about its canonical equivalent. Later in his 1944 paper (2) Post used his concept to tackle the decision problems of recursively enumerable sets of positive integers with interesting success.

However, the full power of this simple definition of language seems not to have been grasped by mathematicians. Undoubtably it can produce deep theorems about languages in general but certainly one of its major applications will be in the study of particular languages. The expression of a language in its canonical form may have no special advantages. If a language is itself inelegantly expressed its canonical form will be inelegant.* But it may bring insights to the study of such a language. In other cases, such as that of the spoken languages, the canonical definition seems uniquely able to handle the language's structure. Not only does it readily provide a mechanical procedure for determining the grammatical correctness of a sentence, it seems to promise deep results in semantics, the problem of determining whether a statement in a spoken language has meaning and how meaningful sentences are constructed.

Here two languages shall be studied in detail in their canonical form, a predicate calculus for number theory and English. Gentzen's scheme has been chosen

*See Rosenbloom (3) for the canonical forms of his languages L_1, L_2', L_X, page 182 for the predicate calculus used here because of its simplicity and for the ease in which it reduces to canonical form, along with an original (to the best of the author's knowledge) modification of Kleene's postulates for number theory which makes arithmetic proofs independent of "Gentzen's cut" and so less dependent upon "guessing". As regards the English language only the grammar problem has been considered and that only for statement sentences, tho the method used is readily applicable to questions or commands. The resulting grammar is quite adequate but by no means perfect. A moderate amount of slave labor and diligence in conjunction with decent research facilities would suffice to make it tight. These examples are enough to illustrate the considerable power of Post's definition of language.

A CANONICAL PREDICATE CALCULUS FOR NUMBER THEORY

Calculi for number theory are well known and we need not explain our purpose in constructing one. But it should be noted that we are mainly interested in the structural and formal problems involved. Indeed the canonical approach forces this attitude upon us since it disallows all such verbalisms as "x is substitutable for y" requiring either that we do without them or that we replace them by some set of productions.

First of all we need a finite set of operational variables, the P's of the generalized production, as specified in our definition. We will need only the following sixteen.

91) α, β any string of ones, possibly null.

(ii) x, y are variables, not null, belonging to the class of assertions determined by production B1.

(iii) a, b, c, d are terms, not null, belonging to the class of assertions determined by the productions B1 thru B4.

(iv) A, B are formulae determined by C1 thru C6.

(v) Γ, Δ, Θ are sequents determined by C1 thru C7.

(vi) S is any sequence of terms derivable from rules B1 thru B4 separated by a comma.

(vii) M, N are any strings whatsoever.

We now define our alphabet and set up the productions which will construct our well formed formulae.

Alphabet

$( ) [ ]\ 1\ '\ /\ ,\ ;\ @\ \vdash\ \supset\ \neg\ \forall\ \cdot\ +\ \vee$

Primitive Assertions

A1 ( ) primitive variable

A2 [/] primitive predicate

Variable Formation

B1 $\dfrac{(\alpha)}{(\alpha 1)}$

Term Formation

B2 $\dfrac{a \quad b}{(a+b)}$ B3 $\dfrac{a \quad b}{(a \cdot b)}$ B4 $\dfrac{a}{(a')}$

Formula Formation

C1 $\dfrac{[\alpha/\beta s] \qquad x\ @[\alpha/\beta s]}{[\alpha/\beta 1 s, x]}$ @ to be read as "not free in"

C2 $\dfrac{[\alpha/\beta s]}{[\alpha 1/\beta s]}$

C3 $\dfrac{\mathbf{A} \quad B}{[\mathbf{A} \supset B]}$ C4 $\dfrac{\mathbf{A}}{[\neg \mathbf{A}]}$ C5 $\dfrac{x \quad \mathbf{A}}{[\forall x \mathbf{A}]}$

C6 $\dfrac{(a/b)\mathbf{A}}{\mathbf{A}}$ S1 $\dfrac{\mathbf{A} \quad a \quad b}{\mathbf{A}(a/b)}$

See substitution rules for meaning of (a/b)

Sequent Formation

C7 $\dfrac{\Gamma \quad \Delta}{\Gamma ; \Delta}$

The concept of a term or variable being free or not free in a formula or term assumes a mechanical procedure for deciding this freedom. Herewith is such a procedure in canonical form.

Free Term Test

@ to be read as "not free in"

D1 $\dfrac{(\alpha)}{(\alpha 1) \;@\; ()}$ D2 $\dfrac{(\alpha)}{() \;@\; (\alpha 1)}$ D3 $\dfrac{(\alpha) \;@\; (\beta)}{(\alpha 1) \;@\; (\beta 1)}$

D4 $\dfrac{x \;@\; a \qquad x \;@\; b}{x \;@\; (a \cdot b)}$ D5 $\dfrac{x \;@\; a \qquad x \;@\; b}{x \;@\; (a \quad b)}$ D6 $\dfrac{x \;@\; a}{x \;@\; (a')}$

D7 $\dfrac{a \qquad b \qquad x}{(a+b) @ x}$

D8 $\dfrac{a \qquad b \qquad x}{(a \cdot b) @ x}$

D9 $\dfrac{a \qquad x}{(a') @ x}$

D10 $\dfrac{c \qquad a @ b \qquad d}{a @ b \vee c @ d}$

D11 $\dfrac{c \qquad a @ b \qquad d}{c @ d \vee a @ b}$

D12 $\dfrac{(a') @ b \qquad (a') @ c}{(a') @ (b \cdot c)}$

D13 $\dfrac{(a') @ b \qquad (a') @ c}{(a') @ (b+c)}$

D14 $\dfrac{(a \cdot b) @ c \qquad (a \cdot b) @ d}{(a \cdot b) @ (c+d)}$

D15 $\dfrac{(a \cdot b) @ c}{(a \cdot b) @ (c')}$

D15 $\dfrac{(a+b) @ c \qquad (a+b) @ d}{(a+b) @ (c \cdot d)}$

D16 $\dfrac{(a+b) @ c}{(a+b) @ (c')}$

D17 $\dfrac{(a') @ b \qquad a @ b \vee b @ a}{(a') @ (b')}$

D18 $\dfrac{(a \cdot b) @ c \qquad (a \cdot b) @ d \qquad a @ c \vee c @ a}{(a \cdot b) @ (c \cdot d)}$

D19 $\dfrac{(a \cdot b) @ c \qquad (a \cdot b) @ d \qquad b @ d \vee d @ b}{(a \cdot b) @ (c \cdot d)}$

D20 $$\frac{(a+b) @ c \qquad (a+b) @ d \qquad a @ c \vee c @ a}{(a+b) @ (c+d)}$$

D21 $$\frac{(a+b) @ c \qquad (a+b) @ d \qquad b @ d \vee d @ b}{(a+b) @ (c+d)}$$

D22 $$\frac{a @ S}{a @ [\alpha // \beta S]}$$

D23 $$\frac{a @ S \qquad a @ b}{a @ S,\, b}$$

D24 $$\frac{a @ \mathbf{A}}{a @ [\neg \mathbf{A}]}$$

D25 $$\frac{a @ \mathbf{A} \qquad a @ B}{a @ [\mathbf{A} \supset B]}$$

D26 $$\frac{a @ \mathbf{A} \qquad x @ a}{a @ [\forall x \mathbf{A}]}$$

D27 $$\frac{[\forall y \mathbf{A}]}{y @ [\forall y \mathbf{A}]}$$

D28 $$\frac{a @ \Gamma \qquad a @ \Delta}{a @ \Gamma ; \Delta}$$

We now need some substitution productions to make it explicitly clear what we mean by "substitutable". Let $\mathbf{A}(a/b)$ be read as "a will replace b in $\mathbf{A}$" and $(a/b)\mathbf{A}$ be read as "a has replaced b in $\mathbf{A}$". No formula containing either $\mathbf{A}(a/b)$ or $(a/b)\mathbf{A}$ shall be considered well formed, tho expressions containing these entities shall certainly be assertions of our system. It is to be noted that the substitution productions given here are valid only for constructive proofs, and that if deductive provability is desirable (for its convenience, not its necessity) alterations would have to be made in many of the productions.

Substitution Rules

S2 $$\frac{M[\mathbf{A} \supset B](a/b)N}{M[\mathbf{A}(a/b) \supset B(a/b)]N}$$

S3 $$\frac{M[(a/b)A \supset (a/b)B]N}{M(a/b)[A \supset B]N}$$

S4 $$\frac{M[\neg \mathbf{A}](a/b)N}{M[\neg \mathbf{A}(a/b)]N}$$

S5 $$\frac{M[\neg (a/b)\mathbf{A}]N}{M(a/b)[\neg \mathbf{A}]N}$$

S6 $$\frac{M[\forall x\mathbf{A}](a/b)N \qquad x \mathbin{@} a \qquad x \mathbin{@} b}{M[\forall x\mathbf{A}(a/b)]N}$$

S7 $$\frac{M[\forall x(a/b)\mathbf{A}]N \qquad x \mathbin{@} a \qquad x \mathbin{@} b}{M(a/b)[\forall x\mathbf{A}]N}$$

S8 $$\frac{M[\forall y\mathbf{A}]N \qquad x \mathbin{@} [\forall y\mathbf{A}]}{M[\forall y\mathbf{A}(x/y)]N}$$

S9 $$\frac{M[\forall y(x/y)\mathbf{A}]N}{M[\forall x\mathbf{A}]N}$$

Rules S8 and S9 are unnecessary but convenient in cases where a bound variable blocks an intended substitution. They can be avoided altogether by the proper choice of bound variables.

S10 $$\frac{M(c')(a/b)N \qquad (c') @ b \vee b @ (c')}{M(c(a/b)')N}$$

S11 $$\frac{M((a/b)c')N}{M(a/b)(c')N}$$

S12 $$\frac{M(c \cdot d)(a/b)N \qquad (c \cdot d) @ b \vee b @ (c \cdot d)}{M(c(a/b) \cdot d(a/b))N}$$

S13 $$\frac{M((a/b)c \cdot (a/b)d)N}{M(a/b)(c \cdot d)N}$$

S14 $$\frac{M(c + d)(a/b)N \qquad (c + d) @ b \vee b @ (c + d)}{M(c(a/b) + d(a/b))N}$$

S15 $$\frac{M((a/b)c + (a/b)d)N}{M(a/b)(c + d)N}$$

S16 $$\frac{M[\alpha/\beta\ S](a/b)N}{M[\alpha/\beta\ S(a/b)]N}$$

S17 $$\frac{M,(a/b)N}{M(a/b),N}$$

S18 $$\frac{M[\alpha/\beta\ (a/b)S]N}{M(a/b)[\alpha/\beta\ S]N}$$

S19 $$\frac{Mx(a/b)N \qquad b @ x}{M(a/b)xN}$$

S20 $$\frac{Mb(a/b)N}{M(a/b)aN}$$

S21 $$\frac{\Gamma \vdash \Delta \qquad a \qquad x}{\Gamma(a/x) \vdash \Delta(a/x)}$$

S22 $$\frac{(a/x)\Gamma \vdash (a/x)\Delta}{\Gamma \vdash \Delta}$$

S23 $$\frac{a \qquad \Gamma \vdash \Delta \quad x \mathbin{@} \Gamma;\Delta}{\Gamma(x/a) \vdash \Delta(x/a)}$$

S24 $$\frac{(x/a)\Gamma \vdash (x/a)\Delta}{\Gamma \vdash \Delta}$$

S25 $$\frac{M;(a/b)N}{M(a/b);N}$$

The rules of inference are straight adaptions from Gentzen's system, which is itself almost in canonical form.

Rules Of Inference

F1 $$\frac{A}{A \vdash A}$$

F2 $$\frac{\Gamma ; A \vdash B ; \Delta}{\Gamma \vdash [A \supset B] ; \Delta}$$

F3 $$\frac{\Gamma \vdash A ; \Delta \qquad \Gamma ; B \vdash \Delta}{\Gamma ; [A \supset B] \vdash \Delta}$$

F4 $$\frac{\Gamma ; A \vdash \Delta}{\Gamma \vdash [\neg A] ; \Delta}$$

F5 $$\frac{\Gamma \vdash A ; \Delta}{\Gamma ; [\neg A] \vdash \Delta}$$

F6 $$\frac{\Gamma \vdash A ; \Delta \qquad x \mathbin{@} \Gamma ; \Delta}{\Gamma \vdash [\forall x A] ; \Delta}$$

F7 $$\frac{\Gamma ; A \vdash \Delta \qquad x \mathbin{@} A \qquad a}{\Gamma ; [\forall x A(x/a)] \vdash \Delta}$$

F8 $$\frac{\Gamma ; A \vdash \Delta \qquad x}{\Gamma ; [\forall x A] \vdash \Delta}$$

S26 $$\frac{\Gamma ; [\forall x (x/a) A] \vdash \Delta}{\Gamma ; [\forall x A] \vdash \Delta}$$

F9 $$\frac{\Gamma \vdash A ; \Delta \qquad \Gamma ; A \vdash \Delta}{\Gamma \vdash \Delta}$$

Structural Rules Of Inference

G1 $$\frac{M\Gamma N \quad A}{M\Gamma ; AN}$$ G2 $$\frac{M\Gamma ; A ; AN}{M\Gamma ; AN}$$ G3 $$\frac{MA ; BN}{MB ; AN}$$

In setting up our productions for number theory it will be convenient to let a = b represent [/11,a,b] and we may arbitrarily choose the symbol () to be the zero of the system and denote it by 0. The following productions are essentially the same as the postulates for number theory which appear in Kleene (4).

Rules For Number Theory

H1 $$\frac{\Gamma ;\ (a\cdot(b')) = ((a\cdot b)+a) \ \vdash \Delta}{\Gamma \vdash \Delta}$$

H2 $$\frac{\Gamma ;\ (a+(b')) = ((a+b)') \ \vdash \Delta}{\Gamma \vdash \Delta}$$

H3 $$\frac{\Gamma ;\ (a+0) = a \quad \vdash \Delta}{\Gamma \vdash \Delta}$$

H4 $$\frac{\Gamma ;\ (a\cdot 0) = 0 \quad \vdash \Delta}{\Gamma \vdash \Delta}$$

H6 $$\frac{\Gamma \ \vdash (a') = 0 \ ; \ \Delta}{\Gamma \vdash \Delta}$$

H7 $$\frac{\Gamma \vdash a = b \; ; \; \Delta}{\Gamma \vdash (a') = (b') \; ; \; \Delta}$$

H8 $$\frac{\Gamma \; ; \; a = b \vdash \Delta}{\Gamma \; ; \; (a') = (b') \vdash \Delta}$$

H9 $$\frac{A \qquad \Gamma \vdash a = b \; ; \; \Delta}{\Gamma \; ; \; A(a/b) \vdash a = b \; ; \; A \; ; \; \Delta}$$

S27 $$\frac{\Gamma \; ; \; (a/b)A \vdash a = b \; ; \; \Delta}{\Gamma \; ; \; A \vdash \Delta}$$

The following productions are for use with the induction procedure. These productions may seem redundantly formulated but there are some inherent difficulties in using an inductive proof constructively. For instance, when we are substituting x for (x') we must first be sure that our formula is free of x's if the procedure is going to be valid. This kind of question does not arise when we are using our induction in the usual manner, ie, when we are trying to decide whether a formula is an assertion or not, rather than trying to construct an assertion. As would be suspected, if we had been formulating our language to be used as a "decision" procedure the productions involved in our induction rule would be much simpler and the "proof tree" of any assertion would be more compact. Constructive proofs in this kind of system are, in general, longer than their corresponding deductive proofs. Unfortunately, in choosing the constructive rather than the deductive approach we have lost much of the "beauty" of our system.

It is also well to note here that once an (a/b) has been introduced into an assertion, all following productions must involve (a/b) until the (a/b) has been eliminated. For example it is not permitted that MA;(a/b)CN $\longrightarrow$ M(a/b)C;AN because the interchange production is good only for well formed formulae and (a/b)C is not well formed by the definition of A and B in G3, which must be assertions derived from the products of productions (C1) thru (C6).

$$\text{S28} \quad \frac{\Gamma \vdash \mathbf{A} \qquad x \mathbin{@} \Gamma;\mathbf{A}}{\Gamma \vdash \mathbf{A}(x/0)} \qquad\qquad \text{S29} \quad \frac{y = x\ ;\Gamma;\ \mathbf{A} \vdash B \qquad x \mathbin{@} \Gamma;B \quad y \mathbin{@} \Gamma;A}{\Gamma\ ;\ A \vdash B((x')/(y'))}$$

$$\text{S30} \quad \frac{\Gamma\ ;\ A\ \vdash ((x')/(y'))B \qquad x \mathbin{@} \Gamma}{\Gamma\ ;\ A\ \vdash\ B(x/(x'))}$$

$$\text{H10} \quad \frac{\Gamma \vdash (x/0)\mathbf{A} \qquad\qquad \Gamma;\ A\ \vdash (x/(x'))\mathbf{A} \qquad\qquad x \mathbin{@} \Gamma}{\Gamma\ \vdash\ \mathbf{A}}$$

In this language proofs of theorems come out in a mechanically direct fashion, whose neatness is perhaps obscured by the poverty of symbols and the unthinking attention given to each small point. Making a few obvious informalisms, however, suffices to show some of the mechanical directness inherent in the language. The proof of such a simple thing as $a=a$ usually covers a messy page as for instance in Kleene, page 84. Here it is done (informally) in a few lines.

$$
\begin{array}{llcll}
 & \text{via F1} \qquad a+0=a & \vdash & a+0=a & \\ \hline
\text{H3} & & \vdash & a+0=a & a=x \\ \hline
\text{H9} & (a=x)(a+0/a) & \vdash & a=x;\ a+0=a & \\ \hline
\text{S} & (a+0/a)(a+0=x) & \vdash & a=x;\ a+0=a & \\ \hline
\text{S27} & a+0=x & \vdash & a=x & \\ \hline
\text{S21} & (a+0=x)(a/x) & \vdash & (a=x)(a/x) & \\ \hline
\text{S} & (a/x)(a+0=a) & \vdash & (a/x)(a=a) & \\ \hline
\text{S22} & a+0=a & \vdash & a=a & \\ \hline
\text{H9} & & \vdash & a=a &
\end{array}
$$

For our proofs we do not need rule F9 (by Gentzen's theorem) but it is a convenient one to have for practical purposes as it allows us to avoid proving over and over again assertions which we know to be theorems; without it, for instance, we would have to prove $a=a$ every time we needed it.

The commutivity of equality is easy to prove.

$$
\begin{array}{l|c}
 & \vdash a=b \qquad\qquad b=x \\ \hline
H9 & (b=x)(a/b) \vdash a=b\ ;\ b=x \\ \hline
S & (a/b)(a=x) \vdash a=b\ ;\ b=x \\ \hline
S27 & a=x \vdash b=x \\ \hline
S21 & (a=x)(a/x) \vdash (b=x)(a/x) \\ \hline
S & a=a \vdash b=a \qquad\qquad a=a \\ \hline
F9 & \vdash b=a
\end{array}
$$

Without the cut we could prove the same thing by using $a+0=a$, and a few more lines.

We could continue in this fashion and prove all of the standard arithmetical theorems, but it is pointless to do so without first making a few more obvious conventions, and perhaps restructuring the productions for deductive analysis, a step which would especially simplify the induction proofs. As it is we have done what we set out to do -- completely formalize the predicate calculus for arithmetic within the framework of Post's canonical definition of language in the simplest garb we could conceive wherein nothing was taken for granted and no procedures were left unwritten to be done by the mathematician's brain. Everything has been successfully reduced to meaningless marks.

A CANONICAL GRAMMAR FOR THE ENGLISH LANGUAGE

In setting up a grammar for a spoken language we are dealing with essentially the same problem as that of the well formed formula of the predicate calculus considered in the previous section by means of the A, B, and C productions. The remaining rules were concerned with procedures for constructing a subclass of well formed formulae, the provable formulae of arithmetic. This indicates that there is no compeling reason for us to stop with the grammar problem. It should be possible to set up productions which would generate meaningful sentences as a subclass of the class of grammatical sentences, meaningful, of course, only with respect to the "reality frame" built into the productions. Such a task promises some formal difficulties and a great deal of clerical work and shall not be attempted here, tho there does not seem to be any essential difference between the two problems other than that one is simpler due to the fact that the elements of the class which we use to define a word are much fewer in the grammatical case.

The alphabet of canonical English consists of the twenty-six letters, the punctuation marks, the space, the numbers, whatever other odd assorted symbols are in current use including intonations and inflections which I shall ignore, and $\vee \wedge \in = .$ Our primitive assertions shall be of three kinds, <u>definitions</u> and primitive strings, the latter divided into <u>primitive words</u> and <u>primitive types</u>. A primitive type can be considered to be a set with base elements consisting of primitive words and all other primitive types which are subsets, including itself; additional elements of the set are constructed by specified grammatical productions. There may be no primitive words in the type set if the grammatical productions only allow phrases as elements. If our object is merely to construct phrase frames or sentence frames we can dilete all primitive words and just consider the strings of types generated by the productions.

We shall let all our definitions be of the following form: $W \approx R$, where W is a list of primitive strings and R a list of primitive types. Each member of W is an element of each member of R, and any primitive type which is a subset of <u>all</u> the R will also appear among the W. The introduction of the complement of a primitive type is possible and has a meaning, but it is such a weak meaning that we shall ignore it.

Other forms of definition are possible depending upon the use to which the system will be put practically. For translation machines an alphabetical listing of all primitive strings in the form x = R would be most convenient, coupled with a transitive law and the convention requiring that no set and its subset appear in the definition of a primitive word, this in order to minimize definition size and maximize availability. We shall not introduce such a convention.

Here R serves as the grammatical definition of each member of W. If we were dealing with semantics, R would actually be the semantic definition, tho of course we would need many more primitive types. The primitive types of semantics would be the so called "undefined terms" which have much the same correspondance as "parts of speech" do to our grammatical types.

Before discussing the simple algebra needed here as a base for the productions, a comment upon the paper of J. Lambek "The Mathematics of Sentence Structure" (5) is in order. At first the ~~author~~ present writer thought that he had a distinctly different system for analyzing grammar. But upon a closer but brief re-examination of Lambek's syntactic calculus it seems that if the procedures herein given are adequate for a complete solution of the grammar problem, then the syntactic calculus (in conjunction with intersection and union productions) is probably also adequate for its solution. This depends upon whether each grammar production can be effectively reduced to a set of what Lambek calls syntactic types.

This can be attempted in this fashion: most grammatical productions are of the form $x \in a \,;\, y \in b \rightarrow xy \in c$ where x and y are strings and a, b and c are primitive types. This can be re-written as $a(a\backslash c) \rightarrow c$ by left division or $(c/b)b \rightarrow c$ by right division. If we then add the syntactic type $(a\backslash c)$ or (c/b) into the definitions and productions where "b" or "a" appears (whether as a primitive type or inside an already present syntactic type) then the production ought to be of no further use providing our algebra can handle right and left division as in the syntactic calculus. By a repeated application of this process it may be possible to eliminate all grammar productions. But the syntactic types can become very complicated. This is not necessarily a drawback. If our definitions become more complicated our algebra becomes simpler, a fact which may be of distinct advantage to a computer. It would be easier for the parsing unit of a translating machine to master the use of the syntactic calculus than it would be for it to master the use of many hundred grammar productions.

The following rules are essentially a very weak boolean algebra, minus among other things a production that would give $x \leftarrow x$ which has been replaced by rule (a1).

W and Y are sequences of primitive strings, and x and y are strings -- necessarily primitive in (a1). S and R are sequences of primitive types. a and b are types -- necessarily primitive in (a1) w and z are strings, possibly empty.

Sentences containing $\wedge$ and $\vee$ may not at first seem meaningful but they have a distinct meaning when we are considering the problem of parsing by machine. Very little of a translating machine's large memory could be in fast access storage. By looking up the definition of each word only once and inserting the intersection of the defining types into the sentence frame this problem is by-passed.

a1 $$\frac{W, x, Y = S, a, R}{x \in a}$$

a2 $$\frac{x \in a \qquad x \in b}{x \in (a \wedge b)}$$

a3 $$\frac{x \in a \qquad y \in b}{x \in (a \vee b)}$$

a4 $$\frac{x \in b \qquad y \in a}{x \in (a \vee b)}$$

a5 $$\frac{w\ x\ z \in a \qquad w\ y\ z \in a}{w\ (x \vee y)\ z \in a}$$

a6 $$\frac{w\ x\ z \in a \qquad y \in c}{w\ (x \wedge y)\ z \in a}$$

a7 $$\frac{w\ y\ z \in a \qquad x \in c}{w\ (x \wedge y)\ z \in a}$$

The system as it stands in rules (a1) thru (a7) trivially has a decision procedure. Later we will show that the system as a whole, with grammar productions, has a decision procedure and so can be used for parsing.

It is amusing to mention the Russell paradox here. The words are all sets which do not contain themselves by definition. Only the primitive types contain themselves. We might ask the question, "Is word a word or a primitive type?" We can of course do no such thing until we can construct such a set. Consider the productions (minus formal details that could be supplied.):

R1 $$\frac{U = \text{word} \qquad y \in x \qquad \text{decide-not } (y \in y)}{U,\ y \in \text{word}}$$

R2 $$\frac{P = \text{prim-type} \qquad y \in x \qquad \text{decide } (y \in y)}{P,\ y \in \text{prim-type}}$$

whence it becomes immediately evident that we have introduced a paradox. (Note: <u>nugo</u> is a singular noun that takes an article.)

$$(R1)\;\frac{U \Rrightarrow \text{word} \qquad \dfrac{W,\ \text{word},\ Y \Rrightarrow S,\ \text{nugo},\ R}{\text{word} \in \text{nugo}}\,(al) \qquad \text{decide-not } (\text{word} \in \text{word})}{\dfrac{U,\ \text{word} \Rrightarrow \text{word}}{\text{word} \in \text{word}}\,(al)}$$

Here we are using our clear cut decision procedure <u>decide-not (word $\in$ word)</u> to assert (<u>word $\in$ word</u>). If we resort to reduction to absurdity and then try to assert (<u>word $\in$ prim-type</u>) we have no better luck.

$$(R2)\;\frac{P \Rrightarrow \text{prim-type} \qquad \dfrac{W,\ \text{word},\ Y \Rrightarrow P,\ \text{nugo},\ R}{\text{word} \in \text{nugo}}\,(al) \qquad \text{decide } (\text{word} \in \text{word})}{\dfrac{P,\ \text{word} \Rrightarrow \text{prim-type}}{\text{word} \in \text{prim-type}}\,(al)}$$

We need the tree <u>decide (word $\in$ word)</u> to assert that (<u>word $\in$ prim-type</u>) but this we cannot do without also asserting <u>decide-not (word $\in$ word)</u>. Thus the paradox. The artificiality of such paradoxes becomes clearly evident in this example. It is eliminated in the usual manner by lifting to a different order the sets by which we discuss our sets. In this case we must replace the sign $\Rrightarrow$ in (R1) and (R2) by, say, $\equiv$ and introduce a new production like (al) replacing $\Rrightarrow$ with $\overset{\sim}{\equiv}$ and $\in$ with, say, $\lessdot$. Then (R1) simply gives us (<u>word $\lessdot$ word</u>) and we cannot assert (<u>word $\lessdot$ prim-type</u>). This is as our reason would demand.

In passing, some of the handicaps under which the grammar productions here given have been obtained should be mentioned. I worked first with my boy's baby dictionary, then graduated to sentences from novels and the Scientific American which I stripped down to manageable form, then graduated to full sized sentences. Working in this way is enormously difficult. For instance it took me a whole day to collect enough sentences representative of which, who, whom, that to make an adequate analysis of that form. "The Structure of English" by Fries (6) was of an immense help, and tho his word classes and groups do not always stand up, they provide some very real meat. What is needed for this type of research is a library of words on IBM cards, sorted alphabetically by docile graduate students, each card punched with the word and having written upon it the sentence in which the word appeared. There also should be a library for words with endings, listed like a rhyming dictionary. With twenty or thirty thousand sentences from representative sources thus filed away, it would be a simple matter to reach in the file, select all the sentences containing the wanted words, say which or much or all, examine them and write out all the productions in which those words appear, compare them with existing productions to see whether the word belongs to any previously existing type and what are its various multiordinal grammatical meanings, and to see whether there are any flaws in the productions. A very tight grammar could be built up in a very short time this way because there can't be more than a few hundred key structural words to examine.

In spite of the difficulties, the grammar herein outlined is fairly adequate, more so than any other grammar the author has seen. "Noun" phrases, "verb" forms, and many of the common modifiers are handled, along with some of the more frequent sentence compounders.

As a small test the author analyzed twenty five sentences averaging thirty words per sentence in sequence from an article in the Scientific American and found only eleven renegade structures, three of these of rare predicate types deliberately ignored, the other eight being modifier phrases (which were not deeply studied -- there are very many) such as "when fresh", "said Cannon", "in short", "so far as possible". Of course, the value of a mechanical grammar depends largely upon its inability to construct ungrammatical forms and this does not show up in parsing if the assertions over-cover the genuine grammatical frames. It would be a simple matter to construct a grammar which would parse with 100% efficiency, if only because it considered all strings of words to be assertions. Still the methods used for constructing grammar rules in this paper clearly indicate the power of Post's canonical definition of language.

A problem which has been completely ignored here is that of bracketing. In all the productions brackets have been inserted, but this is not terribly meaningful as brackets do not exist in a spoken or written language, except in the rudimentary form of pauses, inflections and commas. This of course leads to the fact that not all sentences have unique sentence frames, nor are sentence frames themselves always unique.

There is probably a grammar of brackets implicit in the sentence construction in English, especially spoken English, which eliminates most (but not all) grammatical ambiguities. Certain words seem to carry specific <u>bracket</u> types with themselves. For instance, at the beginning of a modifier, <u>of</u> attaches itself to the element immediately on the left, ie, we always get:

on ((the airplane) in [(the sky) of (America)])

rather than

on ([(the airplane) in (the sky)] of (America))

If we wanted to indicate that the airplane belonged to America we would say, "in the sky on the airplane of America." It should be possible to develop a grammar of brackets.

An appendix of definitions, included at the back, may need to be glanced at by the reader from time to time as the frequency of the introduction of new primitive types comes pretty fast hereafter, and their names are not familiar. We are going to use the constructive approach and build up the basic units of our grammar one by one until we have what we need to assemble sentences. The standard grammatical terms break down completely and we shall not use them. This shift is further necessitated by the fact that we want "parts of speech" which are self reflexive.

First we need to consider the basic predicate frame of English:

$$a \underline{\quad 1 \quad} {}_{b}\text{Predicate}\, {}_{c} \underline{\quad 2 \quad} {}_{d} \underline{\quad 3 \quad}\, e$$

where any or all variable slots may be used depending upon whether our predicate is a function of one, two, or three variables. The latter are rare. The variables are of types which we shall call *nuon*, *pron*, *rom*, and *bekal*, and *vilp*, *valp*. The way in which these variables fill the slots affords a convenient classification of our predicates into six distinct classes. (These classes do not have null intersections) We shall treat only three. Predicate modifiers may occur in positions a, b, c, d, e.

The basic predicate frames are these:

(1) The vil frame.

Slot 1 -- nuon or pron
2 -- empty
3 -- empty

I died
Jesus wept
Satan laughed

(2) The vum frame.

Slot 1 -- nuon or pron
2 -- bekal
3 -- empty

the roasted cannibal tasted delicious
the grapes are sour
democracy seems awful weak

(3) The var frame.

Slot 1 -- nuon or pron
2 -- nuon or rom
3 -- empty

Stalin cherished comrade Trotsky
christians are good honest people
she remembered the misty lake
fate destroyed her casually that afternoon

(4) not treated

Slot 1 -- nuon or pron

2 -- nuon or rom

3 -- nuon

1 2 3 1 2 3

Khrushchev gave his people eggs and butter while Eisenhower presented East Germany lard and flour

1 2 3

they elected me class fool

1 2 3

James appointed him dean

1 2 3

he made his wife toast

(5)

Slot 1 -- nuon or pron

2 -- nuon or rom

3 -- [redacted]vilpor [redacted]valp

1 2 3

the man watched the stripper undress

1 2 3

he made the girl cry

1 2 3

Tarzan made his wife cook monkeys

1 2 3

he saw the axe head bite his toe

(6)

Slot 1 -- nuon or pron

2 -- nuon or rom

3 -- bekal

1 2 3

she made her face tearful

1 2 3

this helpful fact makes analysis possible

1 2 3

it also made the emotions accessible to laboratory measurement

1 2 3

these changes rendered the organism more effective

The command sentences occur with slot 1 empty with the usual fillings for slot 2 and 3, as in: (1) run, fight, goof off; (2) be good; (3) kill it, pass that exam, stop thief; (4) give the boy his money, make your house a home; (5) watch the snake die; (6) make me happy. They are all type [redacted]valp or [redacted]valp and may be taken into account simply by making all elements of [redacted]valp and [redacted]valp also elements of <u>sen</u> (sentence). They introduce no new predicate classifications.

Frame (2) may be inverted as in "mimsy were the borogoves".

Frame (1) may be inverted by an operator like "there" as in "there wept Jesus", "there exist two numbers", "there exists a way".

A fundamental transformation takes place in the basic predicate frame when we operate on predicates from frames (3), (4), and (6) with the elements of the sets <u>tobe</u> or <u>thab</u> (has∨had∨have been). A predicate of three variables is reduced to one of two, a transitive predicate is reduced to an intransitive form. These operations are frequently used in modest scientific papers to facilitate the hiding of the author.

<u> 1 </u> predicate <u> 2 </u> <u> 3 </u>

⟶ <u> 2 </u> (is∨am∨are∨was∨were) predicate <u> 3 </u>

⟶ <u> 2 </u> (have∨has∨had been) predicate <u> 3 </u>

ie, he loved her $\longrightarrow$ she was loved

the girl remembers the lake $\longrightarrow$ the lake has been remembered

she made her face tearful $\longrightarrow$ her face had been made tearful

he gave his people butter $\longrightarrow$ his people were given butter

Naturally no such operation can be executed upon a vil predicate. It is not possible to say "was died". "Worked" in "has been worked" can only be type var. A knowledge of this transformation is important to any translating machine which is never to confuse "operator" with "operated-upon".

A transformation which occurs occassionally is this one: _3_ _1_ predicate _2_ or _2_ _1_ predicate, such as, "Women, I love," "The house which none of us like, I shall destroy," "Pencils, I give you." This form shall not be treated here. It introduces no new predicate types.

Assume that we have a set bekal to draw from. Our nuon set will be divided into singular and plural sets, <u>nuono</u> and <u>nuons</u>. Pron will be divided into singular and plural also, <u>prono</u> containing "he, she, it" and <u>prons</u> containing "I, you, we, they." We will also need two subsets of prons, <u>iprons</u> containing "you, we, they" and <u>i</u> containing "I".

The subsets of our predicate sets attached to the elements of prons, present tense, we will give the appendage -p; those attached to prono, the appendage -ps; those of the past tense -ed; those which appear after "is", the appendage -en; and those forms which end in ing, -ping; ie, the subsets of var are:

varp	varps	vared	varen	varping
lift	lifts	lifted	lifted	lifting
see	sees	saw	seen	seeing

We will assume that these elements have already been premodified if they are to be modified, that is, "already saw" is also an element of vared, "push back" is also an element of varp.

Consider the productions:

b1
$$\frac{x \in \mathrm{vum}\alpha \qquad y \in \mathrm{bekal}}{(x\ y) \in \mathrm{vil}\alpha}$$

seem joyful
sounded very perverse

α ranges over set of predicate endings

b2
$$\frac{x \in \mathrm{avar}\alpha \qquad y \in \mathrm{nuon} \vee \mathrm{rom}}{(x\ y) \in \mathrm{avil}\alpha}$$

remembering his treason
seen the light

b3
$$\frac{x \in \mathrm{aux} \qquad y \in \mathrm{vilp} \vee \mathrm{valp} \vee \mathrm{volp} \vee \mathrm{velp}}{(x\ y) \in \mathrm{auxvil}}$$

could go, might want it
must see the man, will die

b4 $$\frac{x \in \mathrm{do} \qquad y \in \mathrm{Ap}}{(x\ y) \in \mathrm{dAp}}$$

do go, do sound

don't remember

b5 $$\frac{x \in \mathrm{does} \qquad y \in \mathrm{Ap}}{(x\ y) \in \mathrm{dAps}}$$

does see, does like

b6 $$\frac{x \in \mathrm{did} \qquad y \in \mathrm{Ap}}{(x\ y) \in \mathrm{dAed}}$$

did seem, did want

A [redacted] ranges over the set vil, var, vum

b7 $$\frac{x \in \mathrm{dvum}\beta \qquad y \in \mathrm{bekal}}{(x\ y) \in \mathrm{dvil}\beta}$$

did seem good

does sound extreme

β [redacted] ranges over the set of endings p, ps, ed

b8 $$\frac{x \in \mathrm{dvar}\beta \qquad y \in \mathrm{nuon} \vee \mathrm{rom}}{(x\ y) \in \mathrm{dvil}\beta}$$

did remember the man

don't know its name

b9 $$\frac{x \in \mathrm{be} \qquad y \in \mathrm{bekal} \vee \mathrm{nuon} \vee \mathrm{rom} \vee \emptyset}{(x\ y) \in \mathrm{valp}}$$

be good, be human

b10 $$\frac{x \in \mathrm{be} \qquad y \in \mathrm{varen} \vee \mathrm{vilping}}{(x\ y) \in \mathrm{velp}}$$

be sassed, be working

be drinking the water

b11 $$\frac{x \in \text{be} \qquad y \in \text{varping}}{(x\ y) \in \text{beping}}$$

be remembering

be seeing

b12 $$\frac{x \in \text{being} \qquad y \in \text{bekal} \vee \text{varen} \vee \text{nuon} \vee \text{rom}}{(x\ y) \in \text{valping}}$$

being naughty

being taken

being self sufficient

b13 $$\frac{x \in \text{am} \qquad y \in \text{bekal} \vee \text{nuon} \vee \text{rom} \vee \text{varen} \vee \text{vilping} \vee \text{valping}}{(x\ y) \in \text{amil}}$$

am him, am gone

am asking the major

am being a joker

b14 $$\frac{x \in \text{was} \qquad y \in \text{bekal} \vee \text{nuon} \vee \text{rom} \vee \text{varen} \vee \text{vilping} \vee \text{valping}}{(x\ y) \in \text{wasil}}$$

b15 $$\frac{x \in \text{is} \qquad y \in \text{nuon} \vee \text{rom} \vee \text{bekal} \vee \text{varen} \vee \text{vilping} \vee \text{valping}}{(x\ y) \in \text{isil}}$$

b16 $$\frac{x \in \text{were} \vee \text{are} \qquad y \in \text{nuon} \vee \text{rom} \vee \text{bekal} \vee \text{varen} \vee \text{vilping} \vee \text{valping}}{(x\ y) \in \text{waril}}$$

b17 $$\frac{x \in \text{tave}\gamma \qquad y \in \text{vilen}}{(x\ y) \in \text{vil}\gamma}$$

have worked

had struck oil

γ ranges over set of endings ps, ed b18 $$\frac{x \in \text{tavep} \qquad y \in \text{vilen}}{(x\ \ y) \in \text{volp}}$$

have seen the man

(we introduce volp rather than vilp because we do not want "did have seen..."

b19 $$\frac{x \in \text{tave}\beta \qquad y \in \text{been}}{(x\ \ y) \in \text{thab}\beta}$$ *

have been, had just been

b20 $$\frac{x \in \text{thab}\gamma \qquad y \in \text{bekal} \vee \text{nuon} \vee \text{rom} \vee \text{varen} \vee \text{vilping}}{(x\ \ y) \in \text{vil}\gamma}$$

had been good

has been cut

has been cutting wood

b21 $$\frac{x \in \text{thabp} \qquad y \in \text{bekal} \vee \text{nuon} \vee \text{rom} \vee \text{varen} \vee \text{vilping}}{(x\ \ y) \in \text{volp}}$$

have been shot

c1 $$\frac{x \in \text{prons} \vee \text{nuons} \qquad y \in \text{vilp} \vee \text{auxvil} \vee \text{dvilp} \vee \text{volp} \vee \text{viled} \vee \text{dviled}}{(x\ \ y) \in \text{sen}}$$

c2 $$\frac{x \in \text{prono} \vee \text{nuono} \qquad y \in \text{vilps} \vee \text{auxvil} \vee \text{dvilps} \vee \text{wasil} \vee \text{isil} \vee \text{viled} \vee \text{dviled}}{(x\ \ y) \in \text{sen}}$$

c3 $$\frac{x \in \text{i} \qquad y \in \text{amil} \vee \text{wasil}}{(x\ \ y) \in \text{sen}}$$

c4 $$\frac{x \in \text{iprons} \vee \text{nuons} \qquad y \in \text{waril}}{(x\ \ y) \in \text{sen}}$$

* See note in appendix under "thab".

Except for frames (4), (5), (6) which have been neglected and some of the inversions which pop up, the b and c productions cover the predicate structure of the basic statement sentence of English pretty well, tho it might be mentioned that in many people's speech (including the author's own) "got" and "been" have structures not covered here. Fries includes "have to, has to, had to" as a special auxiliary form in itself, but this is not necessary. The infinitive (type invil) serves often as a substantive and fills slot 2 just like any other nuon element. The var predicate <u>have</u> merely has a special affinity for invil elements.

There are certain predicate-like structures involved in many sentences such as: "The continent <u>discovered by Columbus</u> was bleak and menacing. The soldier remembered <u>getting his shots</u>. She wanted <u>to bake a cake</u>." These we will ignore temporarily while we construct our nuon and bekal elements.

A fair number of "substantive" types are needed for an adequate grammar -- he, she, it, they substitutes; who and which correlation forms; classes of substantives with "hard" bases and with predicate bases; and classes with the other special modification structures which crop up. Here we shall be content with three: <u>nugo</u> a singular substantive which takes an article, <u>fugo</u> a singular substantive which does not take an article, <u>nufs</u> a plural substantive which may or may not take an article. See the appendix for representative lists.

Very commonly in English these types modify each other from the right.

d1 $$\frac{x \in \mathrm{nugo} \vee \mathrm{fugo} \qquad y \in B}{(x\ y) \in B}$$

B ranges over [redacted] set nugo, fugo, nufs

school student
target date
air ocean
space frontier
east coast
eye glasses
army officers
Chalumna River
food particles
rock fishes
tree tops

Excluded by the above production are phrases like men friends, formulae maker. Only irregular plurals can be used in the modifying position. These we will ignore.

Let us introduce a type <u>ajekal</u> (which will be a subset of bekal) which contains all more or less non substantive pre-modifiers of substantives. This is by no means a clear cut class. In "the tree tops" "tree" is not an ajekal (because it will not be generally substitutable in ajekal productions) while in "the red coat" "red" can either be type nugo, fugo or ajekal. In "the intuitive man" "intuitive" can only be type ajekal.

Consider three subsets of ajekal; jek, jeker, jekest. We introduce them because they have distinct properties. Jek can be modified by "more" and "most", jeker and jekest cannot, jeker is involved in productions with the word "than" (not discussed), jekest elements can be used as nugo elements, etc.

Also we want to include some elements of varen, vilping, and varping in the ajekal class, but we cannot do this as they stand since we have included all modified forms in these sets. Let avum, avar, avil be sets of completely unmodified predicates, subsets respectively of vum, var, vil. Then avaren, avilping, and avarping are subsets of jek.

d2 $$\frac{x \in \text{nugo} \vee \text{fugo} \qquad y \in \text{avaren} \vee \text{avarping}}{(x\ y) \in \text{ajekal}}$$

A (man made) moon

a (nigger hating) Klaner

their (land invading) offspring

the (spore producing) mushroom

an (instrument carrying) platform

a (Russian designed) spaceship

(potato growing) areas

a (communist inspired) counter revolution

There are many odd ball ajekal elements for which productions shall not be designed here. We have modified-substantives which become ajekal members as: "good size" becomes "good sized", "black hat" becomes "black hatted", "long wind" becomes "long winded". We have jek elements modified on the left by nugo or fugo elements which become ajekal elements as: "world wide", "house high", "age old". We have nugo or fugo elements modified on the left by jek elements which become jek elements as: "odd ball (elements)", "warm climate (varieties)". There are others.

Some of the more common productions involved in the creation of ajekal elements follow, but these should not be taken too seriously as the author had a nowheres near large enough sample of nuon elements to form an adequate classification of the types or productions involved in ajekal construction.

d3 $$\frac{x \in \text{jekly} \qquad y \in \text{jek}}{(x\ y) \in \text{jek}}$$

An (equally striking) finding

d4 $$\frac{x \in \text{less} \vee \text{more} \qquad y \in \text{jek}}{(x\ y) \in \text{jeker}}$$

A (less enjoyable) day

A (more abandoned) woman

d5 $$\frac{x \in \text{too} \qquad y \in \text{jek}}{(x\ y) \in \text{seker}}$$

A (too soft) material

d6 $$\frac{x \in \text{least} \vee \text{most} \qquad y \in \text{jek}}{(x\ y) \in \text{jekest}}$$

The (most characteristic) feature

d7 $$\frac{x \in \text{modek} \qquad y \in \text{jek} \vee \text{jeker} \vee \text{seker} \vee \text{jekest}}{(x\ y) \in \text{ajekal}}$$

This (very remarkable) chain

A (fairly good) house

A (real gone) guy

d8 $$\frac{x \in \text{mucal} \qquad y \in C}{(x\ y) \in C}$$

A (far better) man

A (much too noisy) gigolo

C is type jeker or seker

d9 $$\frac{x \in \text{moda} \qquad y \in \text{ajekal}}{(x\ y) \in \text{ajekal}}$$

A (rather nasty) joke

A (somehow essentially depraved) person

We are now in a position to left-modify our substantives.

d10 $$\frac{x \in \text{ajekal} \qquad y \in \text{nugo}}{(x\ y) \in \text{nago}}$$

d11 $$\frac{x \in \text{ajekal} \qquad y \in \text{fugo}}{(x\ y) \in \text{fago}}$$

d12 $$\frac{x \in \text{ajekal} \qquad y \in \text{nufs}}{(x\ y) \in \text{nafs}}$$

Notice that in rules (d10) to (d12) we have introduced types nago, fago, and nafs, specifically so that the products of these rules cannot be used as input for rule (d1). Thus we avoid expressions like "eye dirty glasses" with "dirty" modifying "glasses".

The article productions below are over simplified. For instance "the other" serves as an article and we make no allowance for this. It should be mentioned that article construction in English is more complicated than is commonly thought. It is also sometimes even ambiguous -- does "most famous men" mean "(most famous) men" or "most (famous men)"? In the latter phrase "most" serves as an article. Also we have treated possessive phrases as articles since they function this way.

e1 $$\frac{x \in \text{ajekal} \qquad y \in D}{(x\ y) \in D}$$

D is the set nug's, fug's, nufs'

e2 $$\frac{x \in \text{pos} \qquad y \in \text{nug's} \vee \text{fug's} \vee \text{nufs'}}{(x\ y) \in \text{eths} \wedge \text{eth}}$$

my men's, our hat's

e3 $$\frac{x \in \text{eths} \qquad y \in \text{nufs'}}{(x\ y) \in \text{eths} \wedge \text{eth}}$$

those men's

e4 $$\frac{x \in \text{eth} \qquad y \in \text{nug's}}{(x\ y) \in \text{eths} \wedge \text{eth}}$$

this man's

e5 $$\frac{x \in \text{fug's}}{(x\ y) \in \text{eths} \wedge \text{eth}}$$

e6 $$\frac{x \in \text{alb} \qquad y \in \text{reths}}{(x\ y) \in \text{eths}}$$

all the, both those

e7 $$\frac{x \in \text{eth} \qquad y \in \text{nago}}{(x\ y) \in \text{fago}}$$

another man

the house

e8 $$\frac{x \in \text{eths} \qquad y \in \text{nafs}}{(x\ y) \in \text{fags}}$$

most men

many enchanting places

Note that nafs is also a subset of fags since no plural need take an article.

Vilping and valping we include as subsets of fago since they are frequently used this way.

Fag elements are modified on the right by a variety of structures, the main right modifier being of the type "in the moon", "of the manor" which we call type nal. To construct these we need a new set frep which vaguely includes those words which are generally called prepositions.

f1 $$\frac{x \in \text{frep} \qquad y \in \text{nuon vrom}}{(x\ y) \in \text{nal}}$$

beneath the hatchet
of these experiments
over the whole envelope
on coated fabrics
for housing radar stations
in the U.S.
under (the direction of Byrd)
above the atmospheric pressure
with seven heads

While we are at it we may as well construct a similar set of predicate modifiers, val, by using the set vrep. Vrep is not identical with frep; neither, unfortunately, is their intersection null.

f2 $$\frac{x \in \text{vrep} \qquad y \in \text{nuon vrom}}{(x\ y) \in \text{val}}$$

toward the mountain
to a true arithmetical formula

by liquidating their officers
as (an invitation to despair)
on (finding the best materials)
to Thompson
from Pierce
in a competitive postwar world
for a few months.
along the highway
thru space
about its day's business
with an electronic machine

There are other words and phrases which seem to belong to type val. "Seaward", "forward", "this Monday", "today", "down-stream", "then", "that way". Being already overburdened we shall not try to set up other productions to generate val phrases such as "this Monday", "that way".

Another defect should be mentioned which is obscured by the fact that val and nal have a large intersection. Altho the val phrases are essentially predicate modifiers they do modify some fag elements, namely those fag elements whose base is essentially a predicate form. For instance we can modify the vilp element "transfer the material" with the val element "to any foreign country". And tho it is not grammatical* to state "((the man) to any foreign country)" we can make the grammatical construct "(the transfer (of the material)(to any foreign country))". Consider also "his return to Washington". This defect can be erased by the introduction of a subset of our substantives composed of the predicate based substantives such as "transfer" "return", or simply by the merger of type nal and val. The latter is not a desirable solution since it introduces a large number of assertions into our system which are patently non-grammatical. We will ignore the defect.

f3 $$\frac{x \in \mathrm{fag}\delta \qquad y \in \mathrm{nal}}{(x\ y) \in \mathrm{fag}\delta}$$

δ [redacted] ranges over set of endings o, s.

Let fagδ be a subset of nuonδ. Then, given a set of right nuon modifiers, we will have constructed about as complex a nuon element as any writer ever dared use. These are the -ping, -en, infinitive, which, and when type modifiers.

e1 $$\frac{x \in \mathrm{vilp} \vee \mathrm{valp} \vee \mathrm{volp} \vee \mathrm{velp}}{(\mathrm{to}\ x) \in \mathrm{invil}}$$

to love

to have been driven [redacted]

---

* that is, not grammatical outside of the Vermont hills where it is acceptable to say, "The storekeeper to East Blackford is a good friend to them."

e2
$$\frac{x \in \text{nuon}\,\delta \qquad y \in \text{invil} \lor \text{varen} \lor \text{vilping} \lor \text{valping}}{(x\ \ y) \in \text{nuon}\,\delta}$$

the man of the manor to see

regions of the corona giving forth powerful emissions

the decision to leave a lost girl walking the street

the decision procedure for the predicate calculus discovered by Gentzen

e3
$$\frac{\begin{array}{l} x \in \text{nuons} \qquad\qquad\qquad\qquad\qquad\qquad y \in \text{whoch} \\ z \in \text{viled} \lor \text{dviled} \lor \text{vilp} \lor \text{dvilp} \lor \text{volp} \lor \text{auxvil} \lor \text{waril} \end{array}}{(x\ \ y\ \ z) \in \text{nuons}}$$

e4
$$\frac{\begin{array}{l} x \in \text{nuono} \qquad\qquad\qquad\qquad\qquad\qquad y \in \text{whoch} \\ z \in \text{viled} \lor \text{dviled} \lor \text{vilps} \lor \text{dvilps} \lor \text{auxvil} \lor \text{wasil} \lor \text{isil} \end{array}}{(x\ \ y\ \ z) \in \text{nuono}}$$

standard items which can be mass-produced by batteries of machines

units of his writing which are the objects of the teacher's consideration

non-human creatures who were really alien

the men who enjoyed these gargantuan decorations

We have [redacted] similar productions involving the set <u>whomch</u> (which contains an empty element), however, to construct this modifier we need the elements of a set named <u>nuvar</u>. Nuvar elements require for their generation more than fifteen productions which shall not be listed here. The construction of these productions is routine and shall be left as an exercise for the reader. They are essentially sen frames with an object-like element knocked off at the end.

e5 $$\frac{x \in \text{nuon}\delta \qquad y \in \text{whomch} \qquad z \in \text{nuvar}}{(x\ y\ z) \in \text{nuon}\delta}$$

The woman I loved

the <u>r</u> which we pronounce
but do not write (in colonel)

a piping noise which the
queen bee can make by forcing
air thru her abdominal spiracles

e6 $$\frac{w \in \text{nuon}\delta \qquad x \in \text{whomch} \qquad y \in \text{sen} \qquad z \in \text{vrep}}{(w\ x\ y\ z) \in \text{nuon}\delta}$$

e7 $$\frac{w \in \text{nuon}\delta \qquad x \in \text{vrep} \qquad y \in \text{owhomch} \qquad z \in \text{sen}}{(w\ x\ y\ z) \in \text{nuon}\delta}$$

let <u>owhomch</u> be whomch minus the empty element computers for which it does not yet have a suitable list of components

one of the most important agencies by which the bees communicate

a substance which initially would not be either attractive or repellent to salmon to which they would be conditioned

(its pages provide) a hundred hours of the most exasperating reading which a man could ask for

the subject of partial differentiation which we devote this chapter to

There are other modifiers of this kind such as <u>when varen</u> which we see in phrases like "this relation when analyzed". A special class of nuon is modified by <u>when sen</u> and <u>where sen</u> as in "that time when we killed the bottle" or "the university where lunkhead got his degree". For these, productions could easily be generated.

"What" is an interesting word. Such phrases as <u>what nuvar</u> or <u>what auxvil</u> or <u>what vilps</u> serve as nuon elements. Consider: "You know <u>what would have happened to them</u>." "We are safe on that score for <u>what it is worth</u>" (it is worth what) "<u>What people looked for most</u> was extravagance and huge dimension." (people looked for most what) "We can validly deduce no set of statements of <u>what actually exists</u> from any proposition about <u>what ought to be done</u> or <u>what ought to exist</u>."

That sen also serves in the same way as a nuon element, and frequently appears as such in scientific papers. "We proved that our matrix is bounded in the fundamental domain F." "That the triple tangents of the curve of intersection of (36) and S at O coincide with (37) can easily be demonstrated."

Words like "whose" have their own peculiar properties. "Whose" attaches itself to a sentence by gobbling up the first article, if one exists, and in this form acts as a nuon modifier. "There is no question of importance whose decision is not comprised in the science of man." "It loses the collar and becomes a carrier whose function is to transfer the spermatozoon to an egg." Productions can be made to handle these.

Except for one other type of nuon modifier, which we will mention when we construct our bekal elements, this completes our study of the nuon class. It is very nearly an exhaustive treatment, but not entirely as the following illustrations show. "Those near the poles circle in a shorter time than those traveling around the equator, and one band may drift as much as 200 knots faster than the next." "I remember the story of how he came to the college." "The wafers are made of clay, talc, and barium carbonate called steatite." (call is a predicate of frame (4)) "Sawdust laced with oatmeal makes a much better soil for mushroom farming than the usual mixture of compost and straw."

"At 28 Lavoisier married Marie Anne Pierrette Paulze, the 14 year old daughter of a leading member of the Ferme."

Modification of pron and rom elements is a subject which has been ignored due to lack of data. English sentence structure seems to discourage their modification in the usual nuon sense, providing alternate outlets. We never see, "Beautiful I went to town," but rather something like, "I was beautiful as I went to town." Such modification does occur, tho, as in, "You don't love poor little ole me anymore."

We shall now turn to a discussion of bekal elements. We have constructed most of these already in our ajekal set but not all of them. Once the modifier which lies between the article and the substantive has reached a certain size it tends to be shoved out toward the right. We never say, "The bride innocent walked to the altar," but we do say, "The bride innocent in white gown and virginal smile walked to the altar." Therefore let us construct a subset of bekal which we shall call <u>vekal</u>.

h1 $$\frac{x \in \text{ajekal} \vee \text{vekal} \qquad y \in \text{val} \vee \text{invil}}{(x\ y) \in \text{vekal}}$$

h2 $$\frac{x \in \text{nuon}\,\delta \qquad y \in \text{vekal}}{(x\ y) \in \text{nuon}\,\delta}$$

the dependent passive patient
eager to rely on authority
figures

1000 B-17s high in the sky
over Germany

an active agent distinguishable
from adrenalin

the parts essential for intense
physical effort

a growing technical culture
hard pressed for predicates

The intersection of ajekal and vekal elements is our bekal set, the study of which we now consider complete. We do not pretend that we have constructed all possible bekal elements, however. Consider, "It really can't be <u>that small</u>."

Predicate modification is an obscene subject which must be approached with fear and trepidation. Let's start with some of the more exotic forms. "Jake could <u>lip read</u> at forty feet." "He <u>hand tooled</u> the few parts (which) he couldn't buy." "General Motors <u>mass produced</u> another batch of monster barges in 1959." These forms seem to have been impressed into use by a growing technical culture hard pressed for predicates. Predicates left-modified by nugo elements, such as mass or hand or left, is a form with no flexibility in English and as such is essentially idiomatic. As idioms they require separate productions.

There are a fair number of vil and var elements which behave superficially like vum and frame (6) elements respectively. By superficially we mean that only a very limited class of bekal elements can appear in the slot 2 or 3 position. A fair number of such examples are given by Fries (6) page 135. If we wish only to parse, no trouble arises by treating them as members of vum and frame (6), but constructively this is a poor solution. Here is one of the many areas where there is no clear cut division between grammar and semantics. It is perhaps here that the reader should note that any exception, no matter how wild, can always be handled by some special production.

Consider the examples: "The shutter always <u>bangs shut</u>." "A student of mine by the name of Darling <u>blushed red</u> when I called her name." "The baby <u>broke loose</u> from his playpen." "Your wash <u>comes clean clean clean</u> with new blue Crud." "The burglers <u>forced</u> their door <u>open</u>." "I have just the talent to <u>make good</u> as a bum." "This theorem don't <u>ring true</u> nohow." "She <u>turned black and blue</u> after I showed her what's what." Etc.

An important set of post predicate modifiers, we'll call it <u>glup</u>, resembles the vrep set but is by no means identical with it. It is this set which keeps English teachers whimpering because its members keep turning up at the end of sentences where prepositions are not supposed to turn up. The structural difference between the glup and vrep sets is well illustrated by the sentences, "She turned in the covers," and "They built up the slope." Are "turned" and "built" avil elements modified by the val elements "in the covers", "up the slope", or are they avar elements modified by "in" and "up" with "the covers" and "the slope" in slot 1? That is, do we have a woman making a bed or sleeping in it, do we have a group of people building a house on a slope or building a slope?

Such ambiguities tend to be eliminated in English by various available methods. If the semantic content is not clear we get, "She turned the covers in," (a valid and common solution to a difficult problem in spite of what conventional grammarians say) or "She turned over in the covers." Differences in the glup and vrep sets also eliminate some of the possible ambiguities. "Out" and "in" seem to be similar words but "out" is not a member of vrep. When we say, "They rolled out the barrel," there is no ambiguity. If we want to indicate what happened at the end of the party we say, "They rolled out of the barrel," "out of" serving the same function in vrep as "out" serves in glup.

A glup never modifies an avum, it modifies an avil directly on the right, and may modify an avar directly on the right or directly after the nuon or rom has been added if the latter element is not too long. Left modification of avil by glup in conjunction with an inversion sometimes occurs in kiddies' books, as in, "Up jumped Peter Rabbit," but according to Fries this inversion does not occur in modern spoken English. Glup elements sometimes are used as prefixes as in, "overstay, overeat, upend, upchuck."

11 $$\frac{x \in \mathrm{avar}\alpha \qquad y \in \mathrm{glup}}{(x\ y) \in \mathrm{bvar}\alpha}$$

stamp out, throwing up

take down, work up (courage)

12 $$\frac{x \in \mathrm{avar}\alpha \qquad y \in \mathrm{nuon} \vee \mathrm{rom} \qquad z \in \mathrm{glup}}{(x\ y\ z) \in \mathrm{bvil}\alpha}$$

pick the man up

knock the man down

living the story down

13 $$\frac{x \in \mathrm{avil}\alpha \qquad y \in \mathrm{glup}}{(x\ y) \in \mathrm{bvil}\alpha}$$

get up, gets down, got out

dying down, fell over

The elements of tobe are also modifiable by glup as in, "the sun is up," but our notation for tobe, which isn't the best, makes this difficult. We shall ignore it.

The predicate modifier sets moda and jekly which we have seen before in conjunction with the construction of ajekal elements have versatile modification properties.

14 $$\frac{x \in \text{jekly} \vee \text{moda} \qquad\qquad y \in E}{(x\ y) \in E}$$

E ranges over the set bvard, dvild, auxvil, bvild, valp, velp, volp, valping, amil, wasil, isil, waril, invil

clearly could go

somehow did not want it

immediately went

kindly get out of my house

always sing the same song

shortly following its birth

15 $$\frac{x \in F \qquad\qquad y \in \text{not}}{(x\ y) \in F}$$

F ranges over am, is, was, were, are, do does, did, aux, $\underbrace{\text{have, has, had,}}_{\text{tave}}$

is not, must not, does not

was not, might not, will not

16 $$\frac{x \in G \qquad\qquad y \in \text{jekly} \vee \text{moda} \vee \text{val}}{(x\ y) \in G}$$

G ranges over am, is, was, were, are, do does, did, aux, $\underbrace{\text{have, has, had,}}_{\text{tave}}$ be, been

Sets like val, vilping, varen, infel seldom premodify predicates, and the only modifier which appears with any regularity between the var element and slot 2 is the glup group. Such sentences are understandable but usually don't ring true, ie, "The man to the station went to get his wife," and "She remembered that evening the man."

But the author has met: "Electrons have seen inside it." "It was believed that the young of a species possessed an inborn capacity for following only their own parents." "This machine has, as its insulator, a tiny slab composed of high purity titanates." "The mushroom, is, in essence, a spore bearing and spore distributing structure." ".... which in turn might break up spontaneously by the fission process."

Some more common modification productions are:

17 $$\frac{x \in H \qquad y \in \text{val} \vee \text{invil} \vee \text{moda} \vee \text{jekly}}{(x\ y) \in H}$$

H ranges over the set varen, vilα, volp, valp, varping, amil, wasil, isil, waril

18 $$\frac{x \in \text{joen} \qquad y \in \text{sen}}{(x\ y) \in \text{val}}$$

wherever they went

(no matter where) they did it

whenever I came aboard

when the vegetation is most active

because she waited

after my time was up

before atoms can be studied

19 $$\frac{x \in \text{vilping} \vee \text{varen} \vee \text{infel} \vee \text{val} \vee \text{moda} \vee \text{jekly} \qquad y \in \text{sen}}{(y\ x) \vee (x\ y) \in \text{sen}}$$

This completes our by no means adequate analysis of predicate modification.

We have neglected to mention the type joiners such as "and" and "or", and "either-or", "neither-nor", "not-but", etc. The productions involved are trivial providing we are careful to take into account such small points as the fact that "and" joining two singular substantives gives a plural substantive.

A whole subject completely ignored involves productions which transform one sentence frame into an equivalent frame, or integrates two sentence frames into one. A simple example follows:

j1 $\dfrac{\text{w x y and x z v} \in \text{a}}{\text{w x y and z v} \in \text{a}}$

I wondered and I waited ----
I wondered and waited

j2 $\dfrac{\text{w x y and z y v} \in \text{a}}{\text{w x and zy v} \in \text{a}}$

his blue coat and her wool coat
---- his blue and her wool coat

Ignoring such productions as given by (j1) and (j2) it can easily be demonstrated that our system has a decision procedure. Note first that every production given involving $\wedge$ or $\vee$ can be rewritten as a finite number of productions not involving $\wedge$ or $\vee$. Define the rank of the assertion x $\in$ a as the number of occurances of the space. For each given production, the rank of a datum is less than the rank of the product, therefore in a finite number of ways any assertion can be reduced to a set of assertions of rank zero. Each of these assertions is either possible or impossible by the application of (a1) in a finite manner. Hence a decision procedure.

To illustrate the method, let us parse a few standard English sentences.

Ex 1

(e7) this ∈ eth　　chapter ∈ nago

this chapter ∈ fago

the ∈ eth　　purpose ∈ nago　　of ∈ frep　　this chapter ∈ nuon

the purpose ∈ fago　　of this chapter ∈ nal

is ∈ is　　the purpose of this chapter ∈ nuon

B [Is the purpose of this chapter ∈ isil]

types ∈ nufs

various ∈ ajekal　　utterances ∈ nuon

these ∈ eths　　various types ∈ nafs　　of ∈ frep

these various types ∈ fags　　Of utterances ∈ nal

these various types of utterances ∈ fags

signal ∈ avarp　　these various types of utterances ∈ nuon

D [signal these various types of utterances ∈ vilp]

D [vilp ∈ vilp]

contrastive ∈ ajekal　　differences ∈ nufs

the ∈ eths　　contrastive differences ∈ nafs

the contrastive differences ∈ fags　　that ∈ whoch

C [the contrastive differences that vilp ∈ nuon]

C [nuon ∈ nuon]

of ∈ frep

the ∈ eths　　patterns ∈ nafs

the patterns ∈ fags　　of nuon ∈ nal

the patterns of nuon ∈ fags

describe ∈ avarp　　the patterns of nuon ∈ nuon

Describe the patterns of nuon ∈ vilp

to describe the patterns of nuon ∈ invil

A [to describe the patterns of nuon ∈ nuono]

A [to describe the patterns of nuon ∈ nuono]

B [is the purpose of this chapter ∈ isil]

---

to describe the patterns of the contrastive differences that signal these various types of utterances is the purpose of this chapter ∈ sen

Let us consider the construction of a few more sentences in a less formal manner.

Ex 2 After the virus had been split into its two components, these components could recombine under suitable conditions to form particles which looked like the original virus and displayed its properties.

after the virus had been split into its two components ∈ val

after ∈ joen

the virus ∈ nuono

had been split into its two components ∈ viled

had been ∈ thabed

split into its two components ∈ varen

split ∈ varen

into its two components ∈ val

into ∈ vrep

its two components ∈ nuon (here our inadequate treatment of articles comes out: "its two" ought to be type eths, instead we get its ∈ pos ∈ eths, two ∈ ajekal, components ∈ nufs) these components could recombine under suitable conditions to form particles which looked like the original virus and displayed its properties $\in$ sen

particles which looked like the original virus and displayed its properties $\in$ nuon

particles $\in$ nufs

which $\in$ whoch

looked like the original virus $\in$ viled

looked $\in$ viled

like the original virus $\in$ val

like $\in$ vrep (<u>like</u> is a peculiar word like <u>as</u> and needs study)

the original virus $\in$ nuon

displayed its properties $\in$ viled

displayed $\in$ avared

its properties $\in$ nuon

its $\in$ eths

properties $\in$ nufs

(note: viled and viled $\in$ viled)

these components $\in$ nuons

these $\in$ eths   components $\in$ nafs

could recombine under suitable conditions to form nuon $\in$ auxvil

could $\in$ aux

recombine under suitable conditions to form nuon $\in$ vilp

to form nuon $\in$ invil   form nuon $\in$ vilp   form $\in$ avarp

recombine under suitable conditions $\in$ vilp recombine $\in$ vilp

under suitable conditions $\in$ val

Ex 3 We have a universal pudding composed of certain known ingrediants mixed in certain proportions

we $\in$ prons

certain known ingrediants mixed in certain proportions $\in$ nuon

certain known ingrediants $\in$ nuon

certain -- this word is multi-ordinal; I believe in this context it should be in the article class. As an ajekal it means "sure". Its singular equivalent would be "a certain", just as "another" is the singular equivalent of "other".

known $\in$ avaren $\in$ ajekal ingrediants $\in$ nufs

mixed in certain proportions $\in$ varen

mixed $\in$ varen

in certain proportions $\in$ val in $\in$ vrep certain proportions $\in$ nuon

have a universal pudding composed of nuon $\in$ vilp

have $\in$ avarp

a universal pudding composed of nuon $\in$ nuon

a universal pudding $\in$ nuon a $\in$ eth universal $\in$ ajekal

pudding $\in$ nug

composed of nuon $\in$ varen composed $\in$ varen of nuon $\in$ val

Ex 4   Little study has been given to identifying the characteristics

little study ∈ nuono   little ∈ ajekal   study ∈ fugo

has been given to identifying the characteristics ∈ vilps

has been ∈ thabps

given to identifying the characteristics ∈ varen

given ∈ varen   to identifying the characteristics ∈ val

to ∈ vrep   identifying the characteristics ∈ vilping ∈ nuono

identifying ∈ avarping   the characteristics ∈ nuon

E5   He has had considerable experience designing printed circuits.

Here the grammatical structure is ambiguous and without recourse to semantics there is no help for us. If we add a structural detail, an article, one or the other meaning is selected for us, ie, "He has had <u>many</u> experience designing printed circuits," or "He has had experience designing <u>many</u> printed circuits." As is, "designing" can modify either "experience" or "circuits" and our moronical procedure cannot tell us which. Assume the intended meaning.

designing printed circuits ∈ vilping   designing ∈ avarping

printed ∈ avaren ∈ ajekal   circuits ∈ nufs

considerable experience vilping ∈ nuon

has had nuon ∈ vilps   he ∈ prono   has ∈ taveps   had nuon ∈ vilen

had ∈ avaren   considerable experience ∈ nuon

With this brief demonstration of our method we close our discussion of English as a canonical language. The methods used are open to great improvement but it is encouraging to note that they can be so improved, that we have not reached an impasse. A refining of types and a complex cataloging operation, but no barrier to the solution of the many outstanding grammar problems, is in sight. We should be able to achieve a constructively rigorous set of grammar productions which will carry us to, or past, the boundry line between grammar and semantics.

APPENDIX OF PRIMITIVE TYPES

TYPE	SUBSETS AND SAMPLE ELEMENTS				
VIL	vilp	vilps	viled	vilen	vilping
	work	works	worked	worked	working
	die	dies	died	died [redacted]	dying
	go	goes	went	gone	going

(seemed pretty) (remembers the man) (had to go)
(work in the mine with tools)

the a prefix means the completely unmodified form
the b prefix means the glup modified form
the d prefix is the do, does, did modified form
the aux prefix is the must work, might die, will go form -- see production b3

avilping is a subset of jek and ajekal and bekal
(a going concern, a working man)

VUM	vump	vumps	vumed	vumen	vumping
	become	becomes	became	become	becoming
	remain	remains	remained	remained	remaining
	feel	feels	felt	felt	feeling

the a prefix is as above

VAR	varp	varps	vared	varen	varping
	have	has	had	had	having
	give	gives	gave	given	giving
	love	loves	loved	loved	loving
	cut	cuts	cut	cut	cutting

the a prefix means the completely unmodified form
the b prefix is the glup modified form

avaren and avarping are subsets of jek, ajekal, bekal

AUX	might, must, can, could, will, would, shall, should, may, won't, shan't, cannot.
TODO	do, does, did, don't, didn't, doesn't. (the auxiliary -- not the <u>do</u> which appears under the var listings.)
TOBE	is, am, was, are, were, ain't, been, be (with -en or -ping a predicate structure, otherwise behaves like a vil, var, or vum.)
VALP	be, be brutal, be a man, be the sugar in my coffee. (the pre-infinitive form of tobe in its vil, var, vum usage.)
VELP	be remembered, be swiming, be cut, be cutting a rug. (the pre-infinitive form of tobe in its auxiliary or predicate usage.)
VALPING	being remembered, being cut in the face, being good, being a man. (the -ping form of tobe in either its vil, var, vum or its predicate usage.)
AMIL WASIL ISIL WARIL	any predicate-tail following, respectively, am, was, is, were or are. (see productions B13 to B16)
TAVE	the auxiliary have, has, had. (forms a predicate structure exclusively with vilen elements) "has taken a watch from me," is the kind of structure it assumes before taking its place among the vil.
THAB	the auxiliary have been, has been, had been. (this set is unnecessary and should be eliminated which can be done easily by making all <u>been</u> forms, ie, been working on the railroad, been good, been taken, been stabbed in the back, elements of vilen.)

VOLP have been going, have seen the man, have failed. (the pre-infinitive form of the tave and thab auxiliaries. always begins with <u>have</u> unless premodified. separated from <u>vilp</u> in order to prevent the formation of assertions like, "We did have been rocking the boat.")

INVIL the infinitive; vilp, valp, volp, velp premodified by "to" as in, to have been drunk, to slowly [redacted] turn the lock, to see certain smiles.

invil is a subset of nuono and is frequently used as a singular substantive, much more so than is commonly thought.

NUVAR (the main use of this set is with <u>which</u> and <u>whom</u> or nothing at all, as a nuon modifier. the object or an object like element is missing, which is to be "filled in" by the modified nuon.)

I loved ___ that morning when the rain came
he wanted Charlie to see ____
Pappy did not want to leave ___ where it was

NUGO (a singular bare substantive which takes an article, is subset of [redacted] nago, [redacted])

man, airplane, decision, reason, Canadian, latch cat, house, formula, officer, use, step, connective, attempt, way, result, center, case, outside, diet, work, size, ability, structure, angle, transfer

FUGO (a singular bare substantive which does not take an article, is a subset of [redacted] fago, [redacted], nuono, nuon.)

smog, time, mud, friendship, pride, Russian, mankind, freedom, money, sight, reality, activity, sunset,

nugo and fugo are by no means mutually exclusive sets

NUFS (a plural bare substantive, which may or may not take an article, is a subset of nafs, fags nuono, nuon)

positions, airplanes, modifiers, men, formulae

NAGO FAGO NAFS	(the ajekal modified forms of nugo, fago, nufs, respectively, tho also containing the unmodified forms. negative function, air movement, bluish-gray smog, large group, white flattened bodies, hand drills
NUG'S FUG'S NUFS'	(The possessive forms of nugo, fugo, and nufs respectively.) wife's, wives', man's, men's, Lincoln's, doctor's
I	(contains one element, I, is a subset of prons.)
IPRONS	we, you, they; (is a subset of prons, pron)
PRONS	I, we, you, they; (is a subset of pron)
PRONO	he, she, it; (is a subset of pron)
PRON	I, we, you, they, he, she, it.
ROM	us, you, them, him, her, it.,me.
NUONO NUONS NUON	(the generalized substantive of our grammar; nuon the undifferentiated form, nuono and nuons are the plural and singular subsets respectively. contains nufs, nafs, fugo, fago fags as subsets, also invil, vilping, valping. God, the meaning of the modification structure with a class 3 word as head, the answer to the enigma of the spirals, one of the most challenging questions, this differential rotation, this issue, the two features of the steady state theory that seem to cause the greatest surprise, etc.
POS	my, your, her, his, its, our, their, John's.
RETHS	(is a subset of eths) the, those, these.

ALB	all, both (this both not to be mixed with both-and) (alb is a subset of eths.)
ETHS	many, any, few, two, some, no, most, other, various, more, such, ten, fourteen
ETH	the, a, an, this, that, every, no, any, either, neither, one, some, much, more, another
UND	and, but, or, not, rather than
ITHOR	either-or, neither-nor, both-and, not-but
FAGS	(an article-fixed nafs, subset of nuons)
JEK	good, bright, cheap, hard, bitter, dirty, beastly bookish, brutal, famous, angelic, spectacular, momentary, peaceful, faithless, lifelike, fortunate, wooden, ragged, confident, creative, readable, meddlesome, desolate. (a subset of ajekal, bekal)
JEKER	better, brighter, cheaper, harder, worse, less
SEKER	too soft, too nice, too bookish, too bitter
JEKEST	best, brightest, cheapest, dirtiest, worst, most, least
JEKLY	brightly, bitterly, momentarily, brutally, cheaply very brightly, terribly brutally, rather cheaply
MODEK	very, quite, fairly, terribly, real, awfully, awful, pretty
MUCAL	much, little, far, "great deal"

MODA	rather, even, just, still, really, so, somehow surely, now, hardly, also, always, already, mostly, merely, actually
AJEKAL	(a left modifier of substantives, contains avilping, avarping, avaren, jek, jeker, jekest, seker, is itself a subset of bekal.)
BEKAL	(is associated with vum predicates in slot 2, contains subsets ajekal and vekal, qualitatively modifies substantives.)
VEKAL	(see production h1)
FREP	beneath, of, over, on, for, in, under, above, with. (as the grammar herein stands with minimal substantive differentiation, there is very little difference between this set and vrep)
VREP	toward, to, by, as, on, from, in, for, along, thru, about, with, out of, back of, off (see note above for frep)
NAL	(see production f1)
VAL	(see production f2)
WHOCH	who, that, which (insufficient differentiation of our substantives prevents the breakdown of this set into <u>who</u> and <u>that, which</u>.)
WHOMCH	whom, which, and an empty element. (see note above. associated with set nuvar. owhomch is subset of whomch, minus the empty element.)
GLUP	up, down, out, back, off, over
JOEN	wherever, whenever, when, where, no matter where, because, long after, after, before

REFERENCES

(1) Emil L. Post, Formal Reductions of the General Combinatorial Decision Problem, American Journal of Mathematics, Vol 65, '43, pp 197-215.

(2) ____________, Recursively Enumerable Sets of Positive Integers and their Decision Problems, Bulletin American Math Society, Vol 50, '44, pp 284-315.

(3) Paul C. Rosenbloom, The Elements of Math Logic, New York '50 Chap IV, General Syntax of Language.

(4) Stephen C. Kleene, Introduction to Metamathematics, New York '52,

(5) Joachim Lambek, The Mathematics of Sentence Structure, American Mathematical Monthly, Vol LXV, No. 3, March '58, pp 154-170.

(6) Charles C. Fries, The Structure of English, New York '52.