A computational algebraic approach to English grammar

J. Lambek, McGill University, Montreal

Whereas type-logical grammars treat syntactic derivations as logical proofs, usually represented by two-dimensional diagrams, I here wish to defend the view that people process linguistic information by one-dimensional calculations and will explore an algebraic approach, based on the notion of a "pregroup", a partially ordered monoid in which each element has both a left and a right "adjoint". As a first approximation, say to English, one assigns to each word one or more "syntactic types", elements of the free pregroup generated by a partially ordered set of "basic types", in the expectations that the grammaticality of a string of words can be checked by a calculation on the corresponding types. According to G.A. Miller, there is a limit to the temporary storage capacity of our short-term memory, which cannot hold more than seven (plus or minus two) "chunks" of information at any time. We will here explore the possibility of identifying these chunks with "simple types", which are obtained from basic types by forming iterated adjoints. In a more speculative vein, we will attempt to find out how so-called "constraints on transformations" can be framed in the present algebraic context.

1. Introduction

According to anecdotal tradition, Pythagoras (circa 570 BC) had divided mathematics into four disciplines: arithmetic, geometry, music and astronomy. More probably, this was done by a member of his school, Archytas of Tarentum, a true mathematician, whereas Pythagoras is now believed to have been more of a cult leader and politician.

About a thousand years later, Boëthius (480 - 524 AD) proposed this same *quadrivium* (the four ways) as a prerequisite for the study of philosophy. [1)] For about another thousand years, the quadrivium constituted the advanced undergraduate curriculum at medieval universities in Europe. However, three more elementary subjects were required in preparation, the so-called *trivium*, [2)] consisting of logic, grammar and rhetoric.

It was only in the nineteenth century that logic was accepted as a branch of mathematics (Boole, Peirce, Schroeder,.....) and only in the twentieth that grammar was accepted by certain enthusiasts, in particular, so-called *categorial grammar* (Ajdukiewicz, Bar-Hillel,.....).

In my own early endeavours [1958], I conceived of syntax as a substructural logical system, the "syntactic calculus", which may be briefly described as positive propositional intuitionistic logic minus Gentzen's three structural rules: weakening, contraction and interchange.

Later I abandoned this approach for various reasons. [3)] Even now, I feel that the usual way of presenting proofs is two-dimensional, whereas human speakers and hearers are better equipped to make rapid one-dimensional calculations.

I now favour a one-dimensional algebraic approach to grammar. To prove an equation, a student will try to simplify each side, until the simplified terms will agree in the middle. If she is lucky, she need work only on the left-hand side. I think this is a good analogy for what goes on in language processing.

2. Pregroup grammar.

The algebraic system I have in mind is a partially ordered monoid in which each element a has both a left adjoint a^ℓ and a right adjoint a^r such that

$$a^\ell a \to 1 \to a a^\ell,\ a a^r \to 1 \to a^r a.$$

(I use an arrow to denote the partial order.)

I call such a system a *pregroup*, because it becomes a partially ordered group when the two adjoints coincide.[4)]

It is easy to prove from the defining inequalities of a pregroup that

$$\begin{gathered} 1^\ell = 1 = 1^r, \quad a^{r\ell} = a = a^{\ell r}, \\ (ab)^\ell = b^\ell a^\ell,\ (ab)^r = b^r a^r \end{gathered}$$

and that $a \to b$ implies $b^\ell \to a^\ell$ and $b^r \to a^r$.

To start with, we work with the *free* pregroup generated by a partially ordered set of *basic* types, such as:

$$\pi_k = k - \text{th person} \qquad (k = 1, 2, 3)$$

($k = 2$ represents the modern second person singular, as well as all three persons of the plural)

$\mathbf{s}_j$ = declarative sentence in the j-th tense ($j = 1, 2$)
($j = 1$ for the present and $j = 2$ for the past)
$\mathbf{q}_j$ = question in the j-th tense
$\mathbf{i}$ = infinitive of intransitive verb-phrase
$\mathbf{o}$ = direct object

and so on. Sometimes the feature indicated by the subscript j or k is irrelevant, so we omit it and postulate:

$$\mathbf{s}_i \to \mathbf{s}, \quad \mathbf{q}_j \to \mathbf{q}, \quad \pi_k \to \pi.$$

From the basic types we build *simple* types by taking adjoints or repeated adjoints. 5)
Thus, from the basic type a we obtain simple types

$$\cdots a^{\ell\ell},\ a^{\ell},\ a,\ a^{r},\ a^{rr}, \cdots$$

Compound types, or just *types*, are strings of simple types. They are the elements of the *free* pregroup generated by the set of basic types:

1 is the empty string;

multiplication is juxtaposition;

adjoints are defined inductively:

$$1^{\ell} = 1,\ (\alpha\beta)^{\ell} = \beta^{\ell}\alpha^{\ell},\ (\alpha\beta)^{r} = \beta^{r}\alpha^{r};$$

the *partial order* is defined by the rules

$$\frac{\alpha \to \beta}{\beta^{\ell} \to \alpha^{\ell}}, \quad \frac{\alpha \to \beta}{\beta^{r} \to \alpha^{r}}.$$

There is a *metatheorem* [Lambek 1999]: to show that

$$\alpha_1 \cdots \alpha_m \ \to \ \beta_1 \cdots \beta_n$$

we may assume, without loss in generality, that all *generalized* contractions

$$\beta^{\ell}\alpha \to 1,\ \alpha\beta^{r} \to 1 \quad (\alpha \to \beta)$$

precede all generalized expansions

$$1 \to \alpha\beta^{\ell},\ 1 \to \beta^{r}\alpha \quad (\alpha \to \beta).$$

In particular, when $n = 1$, as will be the case in most intended linguistic applications, only contractions are needed, although expansions may still play a theoretical rôle.

Pregroups may be viewed as realizations of a certain substructural logical system, namely *compact bilinear logic*, just as Boolean algebras may be viewed as realizations of the classical propositional calculus. Wojciech Buszkowski [2002] has shown that the above metatheorem may be interpreted as a cut-elimination theorem for compact bilinear logic. Note, however, that this logic is rather unorthodox, inasmuch as it identifies bilinear analogues of disjunction and conjunction.

3. Calculations with types

It seems evident that, when people generate or analyze strings of words, they perform rapid mental calculations, albeit at a subconscious level. It is the contention of the categorial school of linguistics (Ajdukiewicz, Bar-Hillel, $\cdots$, Moortgat, Oehrle, Morrill, Carpenter, Casadio, $\cdots$) that these calculations are not performed on the words themselves, but on the *types* (aka *categories*) which have been assigned to the words in the mental dictionary. In my present view, these types are the elements of a pregroup, in first approximation of a free pregroup.

We shall assign one or more (compound) types to words in the dictionary and proceed to check whether given strings of words are grammatical sentences by performing a calculation in the pregroup.

To get our point across, we confine attention here mainly to verbs and pronouns; but we also admit some plural nouns, because plurals may occur without determiners, which we don't wish to discuss here.

In the examples below, we only exhibit *successful* calculations. It seems that calculations are not performed in parallel, but that people abort calculations that lead nowhere and backtrack to start all over again.

We proceed to look at some examples. We make use of a dash to indicate a Chomskyan trace, for comparison with the prevailing literature.

he likes her

$$\pi_3\ (\pi_3^r \mathbf{s}_1 \mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{s}_1$$

does he like her ?

$$(\mathbf{q}_1 \mathbf{i}^\ell \pi_3^\ell)\ \pi_3\ (\mathbf{i}\mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{q}_1$$

whom does he like $-$?

$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^\ell)\ (\mathbf{q}_1 \mathbf{i}^\ell \pi_3^\ell)\ \pi_3\ (\mathbf{i}\mathbf{o}^\ell)$$

where we postulate $\hat{\mathbf{o}} \to \mathbf{o}$. [7]

A few comments may be in order.

(1) *does* can also have type $\pi_3^r \mathbf{s}_1 \mathbf{i}^\ell$ in emphatic statements.

(2) I follow the late Inspector Morse in preferring *whom* to *who* in the accusative case.

(3) Note that

$$\mathbf{q}^\ell \mathbf{q}_1 \to \mathbf{q}^\ell \mathbf{q} \to 1,\ \hat{\mathbf{o}}^{\ell\ell}\mathbf{o}^\ell \to \hat{\mathbf{o}}^{\ell\ell}\hat{\mathbf{o}}^\ell \to 1,$$

since $\hat{\mathbf{o}} \to \mathbf{o}$ implies $\mathbf{o}^\ell \to \hat{\mathbf{o}}^\ell$.

(4) In our second example, the "surface structure"

$$(\mathbf{q}_1 \mathbf{i} \pi_3^\ell)\pi_3(\mathbf{i}\mathbf{o}^\ell)\mathbf{o}$$

is replaced by the "deep structure" (apologies to Chomsky)

$$\mathbf{q}_1[\mathbf{i}^\ell[\pi_3^\ell \pi_3]\mathbf{i}][\mathbf{o}^\ell \mathbf{o}].$$

To display the two structures simultaneously, we have adopted the above linkages [6)] under the types, in place of square brackets, as easier on the eye. Actually, the right brackets are redundant and we could have just written

$$(\mathbf{q}_1[\mathbf{i}^\ell[\pi_3^\ell)\pi_3(\mathbf{i}[\mathbf{o}^\ell)\mathbf{o}.$$

However, we shall retain *left* brackets only when a potential contraction must be *blocked* for successful calculation, as will be seen in Section 6. These left brackets play a rôle similar to the modalities of Moortgat [1996] and Morrill [1994], but here they have the mere status of punctuation marks and are not elements of the pregroup.

4. Indirect sentences.

The word *say* requires a sentential complement, say of type σ, hence its infinitive has type $\mathbf{i}\sigma^\ell$ and its inflected forms have type $\pi_k^r\mathbf{s}_j\sigma^\ell$. Among the complements we admit both direct and indirect statements and questions. We introduce a few new basic types:

σ = sentential complement

$\bar{\mathbf{s}}_1$ = indirect statement in j-th tense

$\tilde{\mathbf{q}}_j$ = indirect question in j-th tense

and we postulate:

$$\mathbf{s} \to \sigma,\ \mathbf{q} \to \sigma,\ \bar{\mathbf{s}}_j \to \bar{\mathbf{s}} \to \sigma,\ \tilde{\mathbf{q}}_j \to \tilde{\mathbf{q}} \to \sigma.$$

The notation $\bar{\mathbf{s}}$ agrees with the standard X-bar theory [Jackendoff 1977], but we write $\tilde{\mathbf{q}}$ rather than $\bar{\mathbf{q}}$, because indirect questions are not formed from direct questions (with inversion) but from direct statements (without inversion). Thus, the complementizers *that* and *whether* have types $\bar{\mathbf{s}}\mathbf{s}^\ell$ and $\tilde{\mathbf{q}}\mathbf{s}^\ell$ respectively. Here are some examples:

he says he saw her

$$\pi_3\ (\pi_3^r\mathbf{s}_1\sigma^\ell)\ \pi_3\ (\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{s}_1$$

did he say he saw her ?

$$(\mathbf{q}_2\mathbf{i}^\ell\pi_3^\ell)\ \pi_3\ (\mathbf{i}\sigma^\ell)\ \pi_3\ (\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{q}_2$$

whom did he say he saw − ?

$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^\ell)\ (\mathbf{q}_2\mathbf{i}^\ell\pi_3^\ell)\ \pi_3\ (\mathbf{i}\sigma^\ell)\ \pi_3\ (\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \to\ \mathbf{q}$$

whom did he say that he saw − ?

$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^\ell)(\mathbf{q}_2\mathbf{i}^\ell\pi_3^\ell)\ \pi_3\ (\mathbf{i}\sigma^\ell)\ (\bar{\mathbf{s}}\mathbf{s}^\ell)\ \pi_3\ (\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \to\ \mathbf{q}$$

he knows whom he saw −

$$\pi_3^\ell\ (\pi_3^r\mathbf{s}_1\sigma^\ell)\ (\tilde{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell)\ \pi_3\ (\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \to\ \mathbf{s}_1$$

The word *whom*, introducing an indirect question, here has type $\tilde{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$.

Up to now, we have typed the accusative *whom*, but not the nomination *who*. We are tempted to analyze:

who − *saw her* ?

$$(\mathbf{q}\ \mathbf{s}^\ell\pi_3)\ (\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{q}$$

But this type assignment won't account for

who did he say − *saw her* ?

$$(\mathbf{q}\mathbf{s}^\ell\pi_3)\ (\mathbf{q}_2\mathbf{i}^\ell\pi_3^\ell)\ \pi_3\ (\mathbf{i}\sigma^\ell)\quad(\pi_3^r\mathbf{s}_2\mathbf{o}^\ell)\ \mathbf{o}\ \not\to\ \mathbf{q}$$

We therefore follow a different strategy and look at the pseudo-sentence

* *saw her he* ?

$$(\mathbf{q}_1\hat{\pi}_3^\ell\mathbf{o}^\ell)\ \mathbf{o}\ \pi_3$$

with $\hat{\pi}_3 \to \pi_3$ but $\pi_3 \not\to \hat{\pi}_5$. Perhaps an early form of English, like present day German, might have allowed

saw her the man?

We can now assign the type $\mathbf{q}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell$ to *who* and analyze:

who saw her − ?

$$(\mathbf{q}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell)\ (\mathbf{q}\hat{\pi}_3^\ell\mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{q}$$

who did he say saw her − ?

$$(\mathbf{q}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell)\ (\mathbf{q}_2\mathbf{i}^\ell\pi^\ell)\ \pi_3\ (\mathbf{i}\sigma^\ell)\ (\mathbf{q}\hat{\pi}_3^\ell\mathbf{o}^\ell)\ \mathbf{o}\ \to\ \mathbf{q}$$

The new type assignment explains why the following are ungrammatical:

* *who did he say that saw her* − ?

$$\cdots\ \cdots\ \cdots\ (\mathbf{i}\sigma^\ell)\ (\bar{\mathbf{s}}\mathbf{s}^\ell)\ (\mathbf{q}\hat{\pi}_3^\ell\mathbf{o}^\ell)\ \mathbf{o}\ \not\to\ \mathbf{q}$$

* *who did he say whether saw her* − ?

$$\cdots\ \cdots\ \cdots\quad (\tilde{\mathbf{q}}\mathbf{s}^\ell)\quad (\mathbf{q}\hat{\pi}_3^\ell\mathbf{o}^\ell)$$

* *who works* − *and she rests* ?

$$(\mathbf{q}\hat{\pi}_3^{\ell\ell}\mathbf{q}^\ell)\ (\mathbf{q}_1\pi_3^\ell)\ (\mathbf{s}^r\mathbf{s}\mathbf{s}^\ell)\ \pi_3\ (\pi_3^r\mathbf{s}_1)$$

5. Limitations on the short-term memory.

In 1956, the psychologist George A. Miller published an influential paper with the title "The magical number seven plus or minus two: Some limits on our capacity for processing information." The title almost says it all; but Miller elaborates by asserting that we can hold at most 7 ± 2 "chunks" of information in temporary storage. He does not give a precise definition of the word "chunk", and I would like to explore the possibility that, in the present context,

$$\text{chunk} \;=\; \text{simple type.}$$

Let us look at a few examples.

$$\begin{array}{c} \textit{whom did I say he saw ?} \\ (\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\ \pi_1\ (\mathbf{i}\sigma^{\ell})\ \pi_3\ (\pi_3^r\mathbf{s}_2\mathbf{o}^{\ell}) \end{array}$$

After having calculated the type of the initial string *whom did I say*, we hear the words *he saw* and so briefly hold in temporary storage the compound type of exactly seven chunks:

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\sigma^{\ell}\pi_3\ \pi_3^r\mathbf{s}_1\mathbf{o}^{\ell}\ \rightarrow\ \mathbf{q}.$$

In fact, as long as we stick to the grammar developed so far, it seems difficult to exceed seven chunks in temporary storage. Let us enlarge our grammar somewhat by adopting the following new basic types:

$\bar{\mathbf{i}}$ = complete intransitive infinitive (with *to*)

$\mathbf{p}_j$ = participle in j-th tense

$\mathbf{p}$ = plural noun-phrase

$\mathbf{o}'$ = indirect object.

Now consider the following example, in which Miller's magical number seven is never exceeded:

$$\begin{array}{c} \textit{What did she tell you that he had asked me to give her} - ? \\ (\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})(\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\pi_3(\mathbf{i}\sigma^{\ell}\mathbf{o}^{\ell})\mathbf{o}(\bar{\mathbf{s}}\mathbf{s}^{\ell})\pi_3(\pi^r\mathbf{s}_2\mathbf{p}_2^{\ell})(\mathbf{p}_2\bar{\mathbf{i}}^{\ell}\mathbf{o}^{\ell})\mathbf{o}(\bar{\mathbf{i}}\mathbf{i}^{\ell})(\mathbf{i}\mathbf{o}^{\ell}\mathbf{o}'^{\ell})\mathbf{o}' \rightarrow \mathbf{q} \end{array}$$

Here are successive stages in the calculation with different numbers of chunks in temporary storage:

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\mathbf{q}^{\ell}\ \mathbf{q}_2}\mathbf{i}^{\ell}\pi^{\ell} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell}\pi^{\ell} \qquad \text{6 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell}\underbrace{\pi^{\ell}\ \pi_3} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell} \qquad \text{5 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\mathbf{i}^{\ell}\ \mathbf{i}}\sigma^{\ell}\mathbf{o}^{\ell} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\sigma^{\ell}\mathbf{o}^{\ell} \qquad \text{6 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\sigma^{\ell}\underbrace{\mathbf{o}^{\ell}\ \mathbf{o}} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\sigma^{\ell} \qquad \text{5 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\sigma^{\ell}\ \bar{\mathbf{s}}}\mathbf{s}^{\ell} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell} \qquad \text{5 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\mathbf{s}^{\ell}\underbrace{\pi_3\ \pi^{r}}\mathbf{s}_2}\mathbf{p}_2^{\ell} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell} \qquad \text{7 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\underbrace{\mathbf{p}_2^{\ell}\ \mathbf{p}_2}\bar{\mathbf{i}}^{\ell}\mathbf{o}^{\ell} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{i}}^{\ell}\mathbf{o}^{\ell} \qquad \text{6 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{i}}^{\ell}\underbrace{\mathbf{o}^{\ell}\ \mathbf{o}} \ \rightarrow\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{i}}^{\ell} \qquad \text{5 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\bar{\mathbf{i}}^{\ell}\,\bar{\mathbf{i}}\mathbf{i}^{\ell} \;\rightarrow\; \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell} \qquad \text{5 chunks}$$

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell}\,\mathbf{i}\mathbf{o}^{\ell}\;\mathbf{o}'^{\ell} \;\rightarrow\; \mathbf{q}\mathbf{o}'^{\ell} \qquad \text{5 chunks}$$

$$\mathbf{q}\mathbf{o}'^{\ell}\mathbf{o}' \;\rightarrow\; \mathbf{q} \qquad \text{3 chunks}$$

However, it is quite possible to get more than seven chunks:

$$\textit{what did I say you gave him } -?$$
$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\;(\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\;\pi_1\;(\mathbf{i}\sigma^{\ell})\;\pi_2\;(\pi^{r}\mathbf{s}_2\mathbf{o}^{\ell}\mathbf{o}'^{\ell})\;\mathbf{o}'$$

After having performed the indicated contractions and listening to the next two words, we find eight chunks in temporary storage:

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\sigma^{\ell}\pi_2\;\pi^{r}\mathbf{s}_2\mathbf{o}^{\ell}\mathbf{o}'^{\ell} \rightarrow \mathbf{q}\mathbf{o}'^{\ell}.$$

An indirect question of type $\tilde{\mathbf{q}}$ can be the subject of a sentence, hence we postulate

$$\tilde{\mathbf{q}} \rightarrow \pi_3.$$

Now look at the awkward question:

$$\textit{whom does what he drinks } - \textit{ bother } -?$$
$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\;(\mathbf{q}_1\mathbf{i}^{\ell}\pi_3^{\ell})\;(\tilde{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell})\;\pi_3(\pi_3^{r}\mathbf{s}_1\mathbf{o}^{\ell})\;(\mathbf{i}\mathbf{o}^{\ell})$$

Having processed the first three words and hearing the next two, we must briefly hold

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell}\pi_3\;\pi_3^{r}\mathbf{s}_1\mathbf{o}^{\ell} \;\rightarrow\; \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell},$$

with nine chunks in temporary storage.

Of course, most speakers would rephrase this question more pliably:

whom does it bother what he drinks ?

We shall not analyze this sentence here.

With some effort, I can produce an example that exceeds Miller's upper limit of $7+2$.

$$\textit{I know people whom what he drinks } - \textit{ bothers } -?$$
$$\pi_1\;(\pi_1^{r}\mathbf{s}_1\mathbf{o}^{\ell})\quad[\mathbf{p}\;(\mathbf{p}^{r}\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell})\;(\tilde{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell})\pi_3(\pi_3^{r}\mathbf{s}_1\mathbf{o}^{\ell})(\pi_3^{r}\mathbf{s}_1\mathbf{o}^{\ell})$$

Here the plural *people* of type p could be the object of *know*; in fact we must postulate

$$\mathbf{p} \;\rightarrow\; \mathbf{o}.$$

However, in the present situation, we don't want the sentence

I know people

$$\pi_1\ (\pi_1^r \mathbf{s}_1 \mathbf{o}^\ell)\ \mathbf{p} \ \to\ \mathbf{s}_1 .$$

So we block the contraction $\mathbf{o}^\ell\mathbf{p} \to 1$ with the help of a left square bracket. Instead, we perform the contraction $\mathbf{p}\mathbf{p}^r \to 1$ and calculate that the initial string of four words has type $\mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$. Hearing the next three words (before *bother*), we must temporarily store the compound type

$$\mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell\tilde{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell\pi_3\ \pi_3^r\ \mathbf{s}_1\mathbf{o}^\ell \ \to\ \mathbf{s}_1\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell\tilde{\mathbf{q}}$$

of ten chunks. Again, most speakers would avoid this by saying instead:

I know people whom it bothers what he drinks .

6. The block constraint.

Modern linguists of the Chomskyan school have been much concerned with "constraints on transformations", more recently called "barriers to movement" [Chomsky 1986]. It is my impression that such constraints should not be thought of as restrictions on grammatical rules (in fact, traditional grammars don't mention them), but as obstructions to language processing. How these constraints are expressed will, of course, depend on how the grammar has been formulated. Our present approach employs neither transformations nor movement, but talks about adjoints and blocks. In the following section, I will speculate on the pros and cons of formulating constraints in terms of the present machinery.

In [L2000] I had proposed a constraint, best expressed in terms of the hypothetical identification of our simple types with Miller's chunks:

I. Our short-term memory cannot hold two consecutive chunks, both of which are blocked.

In other words, we cannot process $\alpha[\beta[\gamma$ when $\alpha\beta \to 1$ and $\beta\gamma \to 1$. This will take care of one instance of what linguists call "the coordinate structure constraint", other cases of which will be discussed in Section 7 below. First look at

she loves you and me.

$$\pi_3\ (\pi_3 \mathbf{s}_1\mathbf{o}^\ell)\ [\mathbf{o}\ (\mathbf{o}^r\mathbf{o}\mathbf{o}^\ell)\ \mathbf{o}$$

Since *she loves you* is not a constituent of this sentence, the contraction $\mathbf{o}^\ell\mathbf{o} \to 1$ before *you* is blocked. (I won't discuss here why many people say * *she loves you and I.*)

We also have the question

does she love you and me ?

$$(\mathbf{q}_1\mathbf{i}^\ell\pi_3^\ell)\ \pi_3\ (\mathbf{i}\mathbf{o}^\ell)\ [\mathbf{o}\ (\mathbf{o}^r\mathbf{o}\mathbf{o}^\ell)\ \mathbf{o} \to \mathbf{q}_1$$

but will not accept

$$
\begin{gathered}
{}^{*}\ \textit{whom does she love you and} - ? \\
(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{i}^{\ell}\pi_3^{\ell})\ \pi_3\ (\mathbf{i}[\mathbf{o}^{\ell})[\mathbf{o}\ (\mathbf{o}^{r}\mathbf{o}\mathbf{o}^{\ell}).
\end{gathered}
$$

Although this is of type $\mathbf{q}$, it is ruled out by constraint I. (The first square bracket is needed to prevent the question

whom does she love?

from appearing as a constituent.)

There is a problem however. Constraint I rules out not only

$$
(1)\qquad
\begin{gathered}
{}^{*}\textit{what will he paint pictures or} - ? \\
(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{i}^{\ell}\pi^{\ell})\ \pi_3\ (\mathbf{i}[\mathbf{o}^{\ell})\ [\mathbf{p}\ (\mathbf{o}^{r}\mathbf{o}\mathbf{o}^{\ell})
\end{gathered}
$$

as it should, but also the grammatical

$$
(2)\qquad
\begin{gathered}
\textit{what will he paint pictures of} - ? \\
(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{i}^{\ell}\pi^{\ell})\ \pi_3\ (\mathbf{i}[\mathbf{o}^{\ell})\ [\mathbf{p}\ (\mathbf{p}^{r}\mathbf{p}\mathbf{o}^{\ell})
\end{gathered}
$$

which it should not, because both calculations give rise to $\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}[\mathbf{o}^{\ell}[\mathbf{p}$ at the penultimate stage. However, we can avoid the first block by invoking $\mathbf{o}^{\ell} \to \mathbf{p}^{\ell}$ (since $\mathbf{p} \to \mathbf{o}$) after hearing the word *paint*. Now $\mathbf{p}^{\ell} \not\to \mathbf{o}^{\ell}$ (since $\mathbf{o} \not\to \mathbf{p}$), so the first block is no longer needed. Thus (2) gives rise to

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}^{\ell}[\mathbf{p}\ \mathbf{p}^{r}\mathbf{p}\mathbf{o}^{\ell} \ \to \mathbf{q}\ ,$$

but (1) gives rise to

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}^{\ell}[\mathbf{p}\ \mathbf{p}^{r}\mathbf{o}\mathbf{o}^{\ell} \ \not\to \mathbf{q}\ ,$$

since $\mathbf{p}^{\ell}\mathbf{o} \not\to 1$.

7. Coordinate structures.

In Section 6 we assigned the type $\mathbf{o}^{r}\mathbf{o}\mathbf{o}^{\ell}$ to the conjunctions *and* and *or*. Actually, they can have polymorphic type $x^{r}xx^{\ell}$ for many, but not all, values of x, combining two expressions of type x to form another such. In addition to obvious values of x such as $\mathbf{i}, \mathbf{s}, \mathbf{p}, \cdots$, here are some other permitted values of x accompanied by the corresponding values of x^{r} (x^{ℓ} being similar):

I saw and heard them	$(x = \pi^{r}\mathbf{s}_2\mathbf{o}^{\ell},\ x^{r} = \mathbf{o}\mathbf{s}_2^{r}\pi^{rr})$
I saw you and heard them	$(x = \pi^{r}\mathbf{s}_2,\ x^{r} = \mathbf{s}_2^{r}\pi^{rr})$
whom did I see and hear - ?	$(x = \mathbf{i}\mathbf{o}^{\ell},\ x^{r} = \mathbf{o}\mathbf{i}^{r})$
whom did you see and I hear - ?	$(x = \pi\mathbf{i}\mathbf{o}^{\ell},\ x^{r} = \mathbf{o}\mathbf{i}^{r}\pi^{r})$

However, there are many forbidden values of x. For example, $x = \pi_3$ is ruled out, because *he and she* has type π_2. [8] Also $x = \pi i$ is ruled out, because one does not say

$$^{*}\textit{did he come and I see her.}$$

Hence, without invoking any constraint, we know that

$$^{*}\textit{whom did he come and I see} - ?$$

is not acceptable.

Constraint I of Section 6 applies only when $x = \mathbf{o}$. The usual coordinate structure constraint says "no element can be removed from a coordinate structure". Here we formulate this tentatively as follows:

II. When $x = \mathbf{i}, \mathbf{s}, \mathbf{p}, \cdots$, one cannot process $x^{\ell}[xx^{r}xx^{\ell}$ following a double adjoint.

Thus, we cannot process

$$^{*}\textit{whom did you sleep and hear} - ?$$

where we arrive at

$$\hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell}\ [\mathbf{i}\ \mathbf{i}^{r}\ \mathbf{i}\ \mathbf{i}^{\ell} \ \rightarrow\ \hat{\mathbf{o}}^{\ell\ell}\mathbf{i}^{\ell}$$

at the penultimate stage. Nor can we process

$$^{*}\textit{people whom he came and she saw} -$$
$$\mathbf{p}\ (\mathbf{p}^{r}\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell})\pi_3\ (\pi^{r}[\mathbf{s}_2)\ (\mathbf{s}^{r}\mathbf{s}\mathbf{s}^{\ell})\pi_3\ (\pi_3^{r}\mathbf{s}_2\mathbf{o}^{\ell})$$

where one arrives at an intermediate stage at

$$\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell}\ [\mathbf{s}_2\mathbf{s}^{r}\mathbf{s}\mathbf{s}^{\ell} \rightarrow \mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell}\ [\mathbf{s}\mathbf{s}^{r}\mathbf{s}\mathbf{s}^{\ell} \rightarrow \mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^{\ell};$$

but this is ruled out by Constraint II.

Our formulation of the coordinate structure constraint will have to be modified to account for such constructions as *both* - *and* -, *either* - *or* -, *neither* - *nor* -, when the words *both, either, neither* act like left brackets, but are actual English words. We shall refrain from discussing this problem here.

Here is a final example:

$$\textit{I saw photos of him and [paintings of her]}$$
$$\pi_1\ (\pi_1^{r}\ \mathbf{s}_2\mathbf{o}^{\ell})\ [\mathbf{p}\ (\mathbf{p}^{r}\ [\mathbf{p}\ \mathbf{o}^{\ell})\ \mathbf{o}\ (\mathbf{o}^{r}\mathbf{o}\ \mathbf{o}^{\ell})$$

If we try to turn this into a wh-question

$$^{*}\textit{whom did I see photos of him [and paintings of} -]\ ?$$
$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_2\ \mathbf{i}^{\ell}\pi^{\ell})\ \pi_1\ (\mathbf{i}[\mathbf{o}^{\ell})\ [\mathbf{p}\ (\mathbf{p}^{r}[\mathbf{p}\mathbf{o}^{\ell})\ \mathbf{o}$$

we are led to the intermediate stage

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\ [\mathbf{o}^{\ell}\ [\mathbf{p}$$

and so Constraint I applies. However, we can circumvent the first square bracket by applying $\mathbf{o}^{\ell} \to \mathbf{p}^{\ell}$, since $\mathbf{p}^{\ell} \not\to \hat{\mathbf{o}}^{\ell}$. To proceed, we would have to re-assign the type $\mathbf{p}^{r}\mathbf{p}\mathbf{p}^{\ell}$ to *and*. But then we obtain

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}^{\ell}\ [\mathbf{p}\mathbf{p}^{r}\mathbf{p}\mathbf{p}^{\ell}$$

which violates Constraint II.

8. The double adjoint constraint.

Let us look at the question

should I say who saw her − ?

$$(\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\ \pi_1\ (\mathbf{i}\sigma^{\ell})\ (\tilde{\mathbf{q}}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_2\hat{\pi}_3^{\ell}\mathbf{o}^{\ell})\mathbf{o}$$

and try to form a wh-question from this:

**whom should I say who saw − − ?*

$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\ \pi_1\ (\mathbf{i}\sigma^{\ell})\ (\tilde{\mathbf{q}}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_2\hat{\pi}_3^{\ell}\mathbf{o}^{\ell})$$

Note the two consecutive traces; but this will not serve as an explanation (see below) why this apparently grammatical sentence is not acceptable. At the penultimate stage of the calculation, we obtain

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\sigma^{\ell}\tilde{\mathbf{q}}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell}\ \to\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell},$$

so I will conjecture the following constraint:

III. We cannot process two consecutive double adjoints.

Before passing to other examples illustrating Constraint III, let us check that two consecutive traces are permitted. We look at the passive construction:

he was killed −

$$\pi_3\ (\pi_3^{r}\mathbf{s}_2\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell})\ (\mathbf{p}_2\mathbf{o}^{\ell})\ \to\ \mathbf{s}_2$$

recalling that $\mathbf{o}^{\ell} \to \hat{\mathbf{o}}^{\ell}$. This may be turned into a yes-or-no question

was he killed − ?

$$(\mathbf{q}_2\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell}\pi_3^{\ell})\ \pi_3\ (\mathbf{p}_2\mathbf{o}^{\ell})\ \to\ \mathbf{q}_2$$

However, according to the strategy proposed in Section 4, we should consider the pseudo-question

**was killed (the man) − ?*

$$(\mathbf{q}_2\hat{\pi}_3^{\ell}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell})\ (\mathbf{p}_2\mathbf{o}^{\ell})\ \pi_3\ \not\to\ \mathbf{q}_2$$

where $\pi_3 \not\to \hat{\pi}_3$ before introducing the wh-question

$$\textit{who did I say was killed } - \ - \ ?$$
$$(\mathbf{q}\hat{\pi}_3^{\ell\ell}\mathbf{q}_2^{\ell})\ (\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\ \pi_1\ (\mathbf{i}\sigma^{\ell})\ (\mathbf{q}_2\ \hat{\pi}_3^{\ell}\ \hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell})\ (\mathbf{p}_2\mathbf{o}^{\ell})\ \to\ \mathbf{q}$$

Here $\hat{\pi}_3^{\ell\ell}$ and $\hat{\mathbf{o}}^{\ell\ell}$ never occur in juxtaposition, so Constraint III is not violated, although two consecutive traces do appear.

According to the same strategy, we should invoke the pseudo-question

$$^{*}\textit{like him people } ?$$
$$(\mathbf{q}_1\hat{\pi}_2^{\ell}\mathbf{o}^{\ell})\ \mathbf{o}\quad \mathbf{p}\ \not\to\ \mathbf{q}_1$$

to justify the noun phrase

$$\textit{people who like him } -$$
$$\mathbf{p}\quad (\mathbf{p}^{r}\mathbf{p}\hat{\pi}_2^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\hat{\pi}_2^{\ell}\mathbf{o}^{\ell})\mathbf{o}\ \to\ \mathbf{p}\ .$$

Now look at the unacceptable

$$^{*}\textit{whom did I see people who like } - \ - \ ?$$
$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_2\mathbf{i}^{\ell}\pi^{\ell})\ \pi_1\ (\mathbf{i}\ [\mathbf{o}^{\ell})[\mathbf{p}\ (\mathbf{p}^{r}\mathbf{p}\hat{\pi}_2^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\hat{\pi}_2^{\ell}\mathbf{o}^{\ell})$$

This violates Constraint I, but the double block can be avoided if we replace $\mathbf{o}^{\ell}$ by $\mathbf{p}^{\ell}$. Then, at the penultimate stage, we arrive at

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}^{\ell}\ [\mathbf{p}\mathbf{p}^{r}\mathbf{p}\hat{\pi}_2^{\ell\ell}\mathbf{q}^{\ell}\ \to\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\hat{\pi}_2^{\ell\ell}\mathbf{q}^{\ell}$$

and run against Constraint III.

Let us look at two more examples. First consider

$$\textit{What does she like reading books about } - \ ?$$
$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{i}^{\ell}\pi_3^{\ell})\ \pi_3\ (\mathbf{i}\mathbf{p}_1^{\ell})\ (\mathbf{p}_1\mathbf{o}^{\ell})\quad \mathbf{p}\quad (\mathbf{p}^{r}\mathbf{p}\mathbf{o}^{\ell})$$

Here we obtain at the last stage

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\ [\mathbf{o}^{\ell}\ [\mathbf{p}\mathbf{p}^{r}\mathbf{p}\mathbf{o}^{\ell}$$

which is ruled out by Constraint I. Circumventing this by making use of $\mathbf{o}^{\ell} \to \mathbf{p}^{\ell}$, we obtain:

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}^{\ell}\ [\mathbf{p}\ \mathbf{p}^{r}\mathbf{p}\mathbf{o}^{\ell}\ \to\ \mathbf{q}\ .$$

Secondly, consider

$$^{*}\textit{whom does she like reading books which discuss } - \ - \ ?$$
$$(\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{i}^{\ell}\pi_3^{\ell})\ \pi_3\ (\mathbf{i}\mathbf{p}^{\ell})\ (\mathbf{p}_1\mathbf{o}^{\ell})\ \mathbf{p}\ (\mathbf{p}^{r}\mathbf{p}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\hat{\pi}_3^{\ell}\mathbf{o}^{\ell})$$

Here we obtain at the penultimate stage

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\ [\mathbf{o}^{\ell}\ [\mathbf{p}\mathbf{p}^{r}\mathbf{p}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell}$$

to which Constraint I applies. Again, we can circumvent this by using $\mathbf{o}^{\ell} \to \mathbf{p}^{\ell}$, and obtain:

$$\mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}^{\ell}\ [\mathbf{p}\ \mathbf{p}^{r}\ \mathbf{p}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell}\ \to\ \mathbf{q}\hat{\mathbf{o}}^{\ell\ell}\hat{\pi}_3^{\ell\ell}\mathbf{q}^{\ell}$$

But now Constraint III appears.

9. Concluding remarks.

We have explored a variant of categorial grammar, in which types are assigned to each word in the mental dictionary and grammaticality of sentences is checked by a calculation in an algebraic (or logical) system. Here the types are elements of a pregroup, a partially ordered monoid in which each element has a left and a right adjoint.

Confining attention to a miniscule portion of English grammar, we work in the free pregroup generated by a partially ordered set of basic types. In the free pregroup, types are constructed as strings of *simple* types, which are obtained from the basic types by forming repeated adjoints.

We conjecture that the simple types can be identified with Miller's "chunks" of information and have verified that usually not more than seven simple types need to be held in the short-term memory, although some examples require eight and some more far-fetched examples nine or ten.

I believe that algebraic presentations of linguistic information should be essentially associative and non-commutative. Any occurrence of commutativity or any lack of associativity should be explicitly indicated. To license commutativity locally, we assign multiple types to individual words, particularly to verbs. For example, the Latin *amat* should probably be given six types to allow for all the permutations of *puer puellam amat, puer amat puellam* etc, namely the types $(\mathbf{o}^{r})(\pi_3^{r})\mathbf{s}_1,\ (\pi_1^{r})\mathbf{s}_1(\mathbf{o}^{\ell})$ etc. (The parentheses here indicate that both object and explicit subject can be omitted.) To block associativity we have introduced square brackets, but take advantage of the fact that left brackets suffice for our purpose.

We have explored three formal constraints for acceptability of sentences, hoping that these will replace many of the constraints (on transformations) found in the literature: some instances of the coordinate structure constraint and some of the so-called Ross island constraints. I believe that such constraints should not be thought of as rules of grammar, but as restrictions on our ability to process information. More examples should be investigated, in order to see whether our three constraints always apply and whether they are sufficient. One may also ask whether they can be unified under a single principle.

Finally, it should be admitted that *free* pregroups are not likely to explain all of syntax: not all grammatical rules can be stored in the mental dictionary in the form of type assignments. Some additional grammatical rules seem to be required, thus rendering the pregroup no longer free.

For example, the accusative relative pronoun *whom / which / that* of type $\mathbf{p}^r\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$ can be omitted, for instance in the noun-phrase

$$\begin{array}{llll} \textit{people} & \emptyset & \textit{police} & \textit{control} - \\ \mathbf{p} & (\mathbf{p}^r\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell) & \mathbf{p} & (\pi_2^r\mathbf{s}_1\mathbf{o}^\ell) \ \to \mathbf{p} \end{array}$$

Instead of assigning a type to the empty string $\emptyset$, we may adopt a grammatical rule:

$$\mathbf{pso}^\ell \to \mathbf{p}$$

Although this rule is frequently invoked in English, it can lead to sentences which are difficult to parse. Let the reader try her hand at

police police police police police.

Not only is this an acceptable grammatical sentence, it can be parsed in two distinct ways, one of which even gives rise to a tautology. [9)]

This article, like most of my linguistic research, has been confined to syntax, with proper attention to morphology, but ignores semantics and, a fortiori, pragmatics. Whenever I give a talk on my present approach to grammar, someone in the audience will raise the question "but what about semantics?" I usually reply that I prefer Capulet to Montague semantics; but this joke usually falls flat. [10)]

The original syntactic calculus, or rather its proof theory, had a promising link to Montague semantics. As was pointed out by van Benthem [1990] and elaborated impressively by a number of people, as in the books by Morrill [1994] and Carpenter [1997], one can exploit the Curry-Howard isomorphism to associate a lambda term to each deduction. This assumes implicitly that one has first introduced Gentzen's structural rules and then converted the syntactic calculus into Curry's semantic calculus, namely positive intuitionistic propositional logic.

Unfortunately, the logic of pregroups, namely compact bilinear logic, is not a conservative extension of the syntactic calculus. Yet, van Benthem's program can still be carried out in a less deterministic way. For example, a derivation $a \to bcd^\ell$ may be interpreted as a function from a to $b \times c^d$ or as a function from a to $(b \times c)^d$. [11)]

REFERENCES[12)]

M. Barr, ∗-autonomous categories: once more around the track, Theory and Applications of Categories **6**(1999), 5-24.

W. Buszkowski, Cut elimination for the Lambek calculus with adjoints, in: V.M. Abrusci and C. Casadio (eds), New perspectives in logical and formal linguistics, Proc. 5th Rome Workshop, Bulzoni Editore, Rome 2002.

B. Carpenter, Type-Logical Semantics, A Bradford Book, The MIT Press, Cambridge Mass. 1997.

C. Casadio, Logic for grammar. Developments in linear logic and formal linguistics, Università di Bologna, Bolzoni Editore, Rome 2002.

C. Casadio and J. Lambek, A tale of four grammars, Studia Logica **71**(2002).

N. Chomsky, Barriers, MIT Press, Cambridge Mass 1986.

H.B. Curry, R. Feys and W. Craig, Combinatory Logic, North Holland, Amsterdam 1958.

J.C. Fisher, Do as I say, not as I do, Amer. Math. Monthly **93**(1986), 664.

Z.S. Harris, A cyclic cancellation - automaton for sentence well-formedness, International Computation Centre Bulletin **5**(1966), 69-94.

R. Jackendoff, $\overline{X}$ Syntax: A study of phrase structure, Linguistic Inquiry Monograph Two, The MIT Press, Cambridge Mass, 1977.

J. Lambek, The mathematics of sentence structure, Amer. Math. Monthly **65**(1958), 154-170.

.........., Deductive systems and categories in linguistics, in: H.J. Ohlbach and U. Reyle (eds.), Logic, Language and Reasoning, Essays in honour of Dov Gabbay, Kluwer, Dordrecht 1999.

.........., Type grammars revisited, in: A. Lecomte, F. Lamarche and G. Perrier (eds), Logical Aspects of Computational Linguistics, Springer LNAI 1582 (1999), 1-27.

.........., Pregroups: a new algebraic approach to sentence structure, in: C. Martin-Vide and G. Păun (eds.), Recent topics in mathematical and computational linguistics, Editura Academici Române, Bucharest 2000.

.........., Type grammars as pregroups, Grammars **4**(2001), 21-39.

J.D. McCawly, The syntactic phenomena of English, The University of Chicago Press, Chicago 1988.

G.A. Miller, The magical number seven plus or minus two: some limits on our capacity for processing information, Psychological Review **63** (1956), 81-97.

M. Moortgat, Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus, Foris Publications, Dordrecht 1988.

.........., Multimodal linguistic inference, Journal of Logic, Language and Information **5**(1996), 349-385.

.........., Categorial type logics, in: J. van Benthem and A. ter Meulen (eds.), Handbook of Logic and Language, 93-177, Elsevier, Amsterdam 1997.

G.V. Morrill, Type Logical Grammar: Categorial Logic of Signs, Kluwer, Dordrecht 1994.

W. O'Grady and M. Dobrowolsky, Contemporary Linguistic Analysis: An Introduction, 3rd edition, Copp Clark Ltd, Toronto 1996.

J. van Benthem, The semantics of variety in categorial grammar, in: W. Buszkowski, W. Marciszewski and J. van Benthem (eds.). Categorial Grammar, J. Benjamin, Amsterdam 1990.

FOOTNOTES

[1] I can only conjecture that this was not the reason for his execution by order of the Gothic king Theodoric of Ravenna, whom he had served as chief minister.

[2] Hence our derogatory term "trivial".

[3] I was accused of mathematical imperialism by several graduate students of Linguistics, and I was soon converted to Chomsky's more powerful theories.

[4] It becomes a discrete group if the partial order is the equality relation. For the benefit of mathematicians with an understanding of category theory, a pregroup is a special kind of bicategory in which each one-cell has a left and a right adjoint. In another example of such a bicategory, the zero-cells are division rings, the one-cells are finitely generated bimodules, with tensor product as composition, and the two-cells are bilinear transformations.

[5] It seems that double adjoints suffice in practice. Double left adjoints are needed for English, French, Italian and German. German may also require double right adjoints.

[6] The linkages indicating (generalized) contractions may be considered as degenerate forms of the proofnets fashionable in Linear Logic, although they had previously been used by Zellig Harris [1966] in a different context, as was pointed out by Aravind Joshi.

[7] We might have taken $\hat{\mathbf{o}} = \mathbf{o}$ here, but it is useful to have an inequality instead, to help in contexts not studied in the present article, e.g. to account for

$$\textit{whom did you see yesterday} - ?$$

Here *see* should be given the type $\mathbf{i}\hat{\mathbf{o}}\alpha^{\ell}$, α being the type of the adverb *yesterday*, without admitting

$$^{*}\textit{did I see yesterday him } ?$$

[8] For *or* the story is even more complicated. As noticed by Fisher [1986], the American Heritage Dictionary of the English Language manages to formulate a rule in a statement that breaks it: "When the elements [connected by *or*] do not agree in number, or when one or more of them is a personal pronoun, the verb is governed by the element which is nearer."

[9] Hint: *police* can be a verb meaning "control".

[10] Except in Verona, where I recently stayed at Hotel Capuletti.

[11] An exposition of the interplay between these various systems may be found in Casadio and Lambek [2002]. For mathematicians with a categorical bend I repeat here my frequent observation (see e.g. [1999a]) that the proof theory of the syntactic calculus may be viewed as a residuated (biclosed monoidal) category. This may be converted into a cartesian closed category by introducing appropriate equations. Hence Montague semantics, ignoring intensionality, may be viewed as a functor from a residuated category to a cartesian closed one or even a topos. Both of these concepts are due to Bill Lawvere, in his profound investigation into the foundations of mathematics. The semantics of a pregroup grammar may be viewed not as a functor, but as a profunctor from a compact (non-commutative) star-autonomous category, in the sense of Michael Barr [1999], into a cartesian closed one.

[12] I have listed only those sources which are cited in the text and not all the publications that have influenced me, such as many important books by Noam Chomsky, which have undergone considerable transformations from "Syntactic Structures" to "Government and Binding". I must admit that I haven't mastered the most recent maximalist program, which has become current dogma in many linguistics departments.