Pregroups and Natural Language Processing

Joachim Lambek

A *pregroup* is a partially ordered monoid endowed with two unary operations called left and right adjunction. Pregroups were recently introduced to help with natural language processing, as I illustrate here by looking at small fragments of three modern European languages. As it turns out, the apparently new algebraic concept of a pregroup had been around for some time in recreational mathematics, although not under this name.

Mathematical Linguistics

The Pythagoreans, if not the master himself (about 570 BC), divided mathematics into four disciplines: arithmetic, geometry, music, and astronomy. About a thousand years later, Boëthius (AD 450–524) proposed the same *quadrivium* (= four ways) as a prerequisite for the study of philosophy. Except for a lapse during the Dark Ages, this quadrivium constituted the advanced undergraduate curriculum at European universities for another thousand years. However, three more elementary subjects were required for preparation, the so-called *trivium*: logic, grammar, and rhetoric, making up altogether seven liberal arts. Of the *trivial* subjects, logic was accepted as a branch of mathematics in the nineteenth century (Boole, Peirce, Schroeder, $\cdots$), and finally grammar too was admitted, though reluctantly, in the twentieth. It should not come as a surprise that, over the years, several mathematicians have made contributions to the study of language. For example, Eratosthenes and Wallis published works on grammar,[1] and Grassman is perhaps even more famous for his contributions to philology than to mathematics.

Today there are a number of different approaches to grammar, most prominently that of Noam Chomsky, whose system of generative-transformational grammar has evolved considerably over the years. Mathematicians interested in the subject have largely favoured another approach, which is intended to complement the linguists' insight with a computational component. Originally called "categorial grammar", it was pioneered by K. Ajdukiewicz (who had been influenced by E. Husserl and S. Lesniewski) and later expanded by Bar-Hillel.[2] I myself proposed a system of propositional logic without structural rules, which I called "syntactic calculus". Roughly at the same time, Haskell Curry developed what I prefer to call a "semantic calculus", essentially positive intuitionistic propositional logic. According to the so-called Curry-Howard isomorphism, its proof theory is equivalent to the lambda calculus or combinatory logic (more recently also to Bill Lawvere's cartesian closed categories).

A more sophisticated version of the syntactic calculus (or rather a non-associative variant of it), with additional modalities to license associativity and commutativity when needed, was developed by Michael Moortgat, his disciples and collaborators (see, e.g., [24]). On the other hand, the proof theory of Curry's semantic calculus gave rise to Montague se-

This expository article is an expanded version of [18,19]. The underlying research for linguistic applications had been supported by the Social Sciences and Humanities Research Council of Canada.

[1]While the study of grammar may not have the best reputation nowadays, there was a time, in the Middle Ages, when it was supposed to endow a person with magical power, called "glamour", a word derived from "grammar".

[2]When I first submitted [13], I had written "categorical" in place of "categorial", but the referee [Bar-Hillel] kindly suggested that "my typist had made a mistake".

mantics, and its combinatory presentation was exploited for linguistic purposes by Ed Keenan and E. M. Steedman.

The idea behind the various approaches to syntax is this: assign to each word in the dictionary one or more types (originally called "categories", not to be confused with the categories of Sam Eilenberg and Saunders Mac Lane). These types are terms of a logical system or, equivalently, elements of a freely generated algebraic one. Looking at the string of types associated with a string of words, one should be able to check whether the string of words is a grammatical sentence.

Which algebraic or logical system should be employed for studying grammar? My own view has changed over the years, and I now favour "pregroups", a generalization of partially ordered groups, for the algebraic version. The corresponding logical system has been called "compact bilinear logic". In order not to interrupt the flow of the narrative, I invite the reader interested in the history of the pregroup approach to look at the last section of this article.

Pregroups

A *pregroup* is a partially ordered monoid (a semigroup with unity elements) in which each element a has a *left adjoint* a^ℓ and a *right adjoint* a^r such that

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

Here the arrow is used to denote the partial order. (The reader familiar with category theory will recognize that both the arrow and the notion of adjoint are borrowed from there.)

Of course, every partially ordered group is a pregroup in which $a^\ell = a^r$ is the inverse of a and, conversely, the condition $a^\ell = a^r$ ensures that the underlying monoid of a pregroup is a group. Let me mention only one example of a pregroup which is not a group: the monoid of unbounded order-preserving mappings $\mathbb{Z} \to \mathbb{Z}$ under composition. We shall return to this and a related example of a "left pregroup" in the penultimate section.

Before discussing the intended applications to linguistics, let us look at some mathematical properties of pregroups. For example, adjoints are unique, so we might call a^ℓ *the* left adjoint of a. Indeed, suppose a^* is another left adjoint of a, so that

$$a^* a \to 1 \to aa^*,$$

then

$$a^* = a^* 1 \to a^*(aa^\ell) = (a^* a)a^\ell \to 1a^\ell = a^\ell.$$

Thus $a^* \to a^\ell$, and similarly $a^\ell \to a^*$. Here are some elementary properties of pregroups, the proofs of which will be left to the reader:

$$1^\ell = 1, \; (ab)^\ell = b^\ell a^\ell, \text{ if } a \to b \text{ then } b^\ell \to a^\ell,$$

and similarly for right adjoints. Moreover

$$(a^\ell)^r = a = (a^r)^\ell.$$

Free Pregroups

The linguistic application I have in mind leans primarily on pregroups freely generated by partially ordered sets. We begin with a partially ordered set (poset) of *basic types*, which may differ from one language to another, and which is meant to express certain elementary grammatical concepts and their features. From the basic types one forms *simple types* by repeated adjunction. Thus, a simple type has one of the following forms:

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

where a is a basic type. Finally, we define a (compound) *type* to be a string of simple types. The types form a monoid under concatenation (1 being the empty string), which is easily seen to be partially ordered, provided we stipulate that, for any simple type x,

$$x \to y \text{ implies } y^\ell \to x^\ell,$$

hence

$$x \to y \text{ implies } x^{\ell\ell} \to y^{\ell\ell},$$

and similarly for right adjoints. The partial order may be extended to compound types by stipulating that

$$x \to x' \text{ and } y \to y' \text{ imply } xy \to x'y'.$$

Moreover, the partially ordered monoid of types is seen to be a pregroup, with adjunctions defined inductively thus:

$$1^\ell = 1 = 1^r, \; (xy)^\ell = y^\ell x^\ell, \; (xy)^r = y^r x^r.$$

The resulting pregroup is the *free* pregroup generated by the given poset of basic types. (In the technical language of category theory, this means that we have constructed a functor left adjoint to the so-called "forgetful functor" from pregroups to posets.)

Contractions Suffice

What makes free pregroups particularly amenable for computation is the following observation.

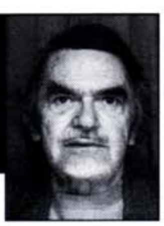
Lemma *When showing that*

$$x_1 \cdots x_m \to y_1 \cdots y_n$$

for simple types x_i and y_j, one may assume without loss of generality that all contractions $a^\ell a \to 1$ and $aa^r \to 1$ precede all expansions $1 \to aa^\ell$ and $1 \to a^r a$.

This was proved in [15], where pregroups were first introduced (under this name) to facilitate language process-

AUTHOR



JOACHIM LAMBEK The author has been on the staff of McGill University since 1946, working now on one branch of mathematics, now on another. He is author of a book on ring theory and co-author of books on categorial logic and on history and philosophy of mathematics. The present article forms part of an ongoing study of non-linear logic and its application to linguistics. He informs *The Intelligencer* (somewhat cryptically: try to make sense of this!) that he has carried out a detailed investigation of the number two.

Department of Mathematics and Statistics
McGill University
Montreal, H3A 2K6
Canada
e-mail: Lambek@math.mcgill.ca

ing.[3] The following argument will give an idea of the proof. A calculation in which an expansion immediately precedes a contraction may look like this: suppose $a \to b$ and $b \to c$, then

$$a = a1 \to ab^r b \to ab^r c \to bb^r c \to 1c = c.$$

This calculation can be replaced by simply citing the transitivity of the partial order

$$a \to b \to c,$$

using neither expansions nor contractions.

Why is this lemma useful to grammarians? Among other things, they are interested in checking whether a given string of words is a grammatically well-formed sentence. To do this, one may look at the corresponding string of simple types, say $x_1 \cdots x_m$, and check whether it reduces to a simple, or even basic, type y, for example the type of a declarative sentence or question. To check whether

$$x_1 \cdots x_m \to y$$

when y is simple, one may assume that all contractions precede all expansions. But, if the right-hand side is simple, there will be no expansions at all! Thus, for sentence verification, we may confine attention to contractions only. This is not to say that expansions are useless; they play a role in proving the mathematical properties of pregroups discussed above.

Some Basic Types

In what follows, we will study the pregroup of types freely generated by a poset of basic types for some very small fragments of three modern European languages. The following common set of basic types will work in all three examples.

π_j = jth personal subject pronoun, where $j = 1, \cdots, 6$ denotes the three persons singular followed by the three persons plural.

Actually, in modern English, the original second person singular has disappeared and has been replaced by the second person plural. Moreover, there is no morphological distinction between the three plural verb forms, hence in English we may put $\pi_2 = \pi_4 = \pi_5 = \pi_6$.

$\mathbf{s}_k$ = declarative sentence in the kth simple tense ($k = 1, 2, \cdots$).

Here $k = 1$ and $k = 2$ stand for the present and past indicative, respectively. English and German also have two subjunctives, but express the future as a compound tense. Literary French has altogether seven simple tenses, but two of these are in the process of disappearing.

$\mathbf{q}_k$ = yes-or-no questions in the kth simple tense,
$\mathbf{o}$ = direct object,
$\mathbf{p}_2$ = past participle of intransitive verb,
$\mathbf{i}$ = infinitive of intransitive verb.

Both of the last-mentioned types may also apply to compound verb phrases. It is convenient to introduce also the types:

π = subject when the person is irrelevant,
$\mathbf{q}$ = yes-or-no question when the tense is irrelevant,
$\bar{\mathbf{q}}$ = question (including yes-or-no questions and wh-questions) and to postulate

$$\pi_j \to \pi, \qquad \mathbf{q}_k \to \mathbf{q} \to \bar{\mathbf{q}}.$$

A Small Fragment of English

We begin by assigning some types to a few English words:

he has type π_3 (= third person subject),
her has type $\mathbf{o}$ (= direct object),
sees has type $\pi_3^r \mathbf{s}_1 \mathbf{o}^\ell$

to indicate that we require a third person subject on the left and a direct object on the right. (This idea may have been anticipated in a chemical analogy by Charles Sanders Peirce [27], who would have said that *sees* resembles a molecule with two unsaturated chemical bonds.)

Now look at the sentence

he sees her
$$\underline{\pi_3 \; (\pi_3^r} \mathbf{s}_1 \underline{\mathbf{o}^\ell) \; \mathbf{o}} \to \mathbf{s}_1$$

We calculate in two steps:

$$\pi_3(\pi_3^r \mathbf{s}_1 \mathbf{o}^\ell) = (\pi_3 \pi_3^r) \mathbf{s}_1 \mathbf{o}^\ell \to 1 \mathbf{s}_1 \mathbf{o}^\ell = \mathbf{s}_1 \mathbf{o}^\ell,$$
$$(\mathbf{s}_1 \mathbf{o}^\ell) \mathbf{o} = \mathbf{s}_1 (\mathbf{o}^\ell \mathbf{o}) \to \mathbf{s}_1 1 = \mathbf{s}_1.$$

It is convenient to indicate contractions by *underlinks*.[4] Similarly, we have

I saw her,
$$\underline{\pi_1 \; (\pi^r} \; \mathbf{s}_2 \underline{\mathbf{o}^\ell) \; \mathbf{o}} \to \mathbf{s}_2,$$

where the first underlink represents the generalized contraction

$$\pi_1 \pi^r \to \pi \pi^r \to 1.$$

In our next example, we make use of two further type assignments:

has has type $\pi_3^r \mathbf{s}_1 \mathbf{p}_2^\ell$,
seen has type $\mathbf{p}_2 \mathbf{o}^\ell$.

The former requires one complement on each side, the latter only a single complement on the right, to give

he has seen her
$$\underline{\pi_3 \; (\pi_3^r} \mathbf{s}_1 \underline{\mathbf{p}_2^\ell)(\mathbf{p}_2} \underline{\mathbf{o}^\ell) \mathbf{o}} \to \mathbf{s}_1.$$

Note in contrast

I have seen her
$$\underline{\pi_1 \; (\pi_1^r} \mathbf{s}_1 \underline{\mathbf{p}_2^\ell)(\mathbf{p}_2} \underline{\mathbf{o}^\ell) \mathbf{o}} \to \mathbf{s}_1$$

you had seen her
$$\underline{\pi_2 \; (\pi^r} \mathbf{s}_2 \underline{\mathbf{p}_2^\ell)(\mathbf{p}_2} \underline{\mathbf{o}^\ell) \mathbf{o}} \to \mathbf{s}_2.$$

Unfortunately, *has* must be assigned a different type in direct questions, namely

$$has : \mathbf{q}_1 \mathbf{p}_2^\ell \pi_3^\ell$$

[3]Buszkowski [5] has shown that this lemma is essentially a *cut elimination theorem* for a certain logical system, here called *compact bilinear logic*.

[4]Such linkages go back to Zellig Harris [10], as I learned from A. K. Joshi.

to obtain

$$\begin{array}{c}\textit{has he seen her?}\\ (\mathbf{q}_1\mathbf{p}_2^{\ell}\pi_3^{\ell})\pi_3(\mathbf{p}_2\mathbf{o}^{\ell})\mathbf{o} \to \mathbf{q}_1\end{array}$$

Not wishing to overload the mental dictionary with multiple type listing, one may adopt certain *metarules*. In the present case, such a metarule would convert the type $\pi_j^r\mathbf{s}_k x^{\ell}$ into $\mathbf{q}_k x^{\ell}\pi_j^{\ell}$. In English, this metarule is restricted to auxiliary verbs, here with $x = \mathbf{p}_2$; but, in German, it applies to all verbs.

I will assign the following type to the object question word

$$\textit{whom} : \bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell},$$

where $\hat{\mathbf{o}} \to \mathbf{o}$. The reader will notice that, following the late Inspector Morse, I distinguish between *whom* and the subject question word *who*, which must be assigned a different type.[5] Thus we have

$$\begin{array}{c}\textit{whom has he seen—?}\\ (\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{p}_2^{\ell}\pi_3^{\ell})\ \pi_3\ (\mathbf{p}_2\mathbf{o}^{\ell})\end{array}$$

since $\mathbf{q}^{\ell}\mathbf{q}_1 \to \mathbf{q}^{\ell}\mathbf{q} \to 1$ and $\hat{\mathbf{o}}^{\ell\ell}\mathbf{o}^{\ell} \to \hat{\mathbf{o}}^{\ell\ell}\hat{\mathbf{o}}^{\ell} \to 1$, the latter in view of the contravariant adjunction. The dash here represents what Chomsky used to call a *trace*. It turns out that double adjoints will always appear in the presence of traces.

It may be of interest to see how the last calculation may be performed step by step when the four words are heard in succession. Here are four stages of the calculation:

whom: $\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell}$

(*whom*) *has*: $(\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})\ (\mathbf{q}_1\mathbf{p}_2^{\ell}\pi_3^{\ell}) \to \bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell}\pi_3^{\ell}$

(*whom has*) *he*: $(\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell}\pi_3^{\ell})\ \pi_3 \to \bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell}$

(*whom has he*) *seen*: $(\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell})\ (\mathbf{p}_2\mathbf{o}^{\ell}) \to \bar{\mathbf{q}}$

When hearing the question, one must hold successively 3, 6, 4, 5, 3, 5, 1 simple types in temporary storage. It is tempting to identify these simple types with G. A. Miller's *chunks of information*. In an influential paper [23], Miller suggested that humans can hold a maximum of seven (plus or minus two) chunks of information in their short-term memory.

The auxiliary verb *be* may be employed for constructing passives in English. In particular, its past participle then has the following type:

$$\textit{been} : \mathbf{p}_2\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2^{\ell}.$$

Notice another double adjoint anticipating a Chomskyan trace in the following example:

$$\begin{array}{c}\textit{she had been seen —}\\ \pi_3\ (\pi^r\mathbf{s}_2\mathbf{p}_2^{\ell})\ (\mathbf{p}_2\hat{\mathbf{o}}^{\ell\ell}\mathbf{p}_2)\ (\mathbf{p}_2\mathbf{o}^{\ell}) \to \mathbf{s}_2.\end{array}$$

Apparenty, most Americans, including some prominent linguists, use *who* in place of *whom*, except when *whom* is governed by a preposition, as in

$$\begin{array}{c}\textit{with whom have I seen her ?}\\ (\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell\ell}\mathbf{q}^{\ell})(\bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell}\mathbf{q}^{\ell})(\mathbf{q}_1\mathbf{p}_2^{\ell}\pi_1^{\ell})\pi_1\ \ (\mathbf{p}_2\mathbf{o}^{\ell})\mathbf{o} \to \bar{\mathbf{q}}.\end{array}$$

Note that this analysis requires the type assignment

$$\textit{with} : \bar{\mathbf{q}}\hat{\mathbf{o}}^{\ell\ell\ell}\mathbf{q}^{\ell},$$

so far the only example of a triple left adjoint. However, people who normally avoid *whom* would probably avoid this construction and replace it by

whom have I seen her with—?

which I will not analyze here further (see [20]).

A Very Short Glance at French

In this section, which Francophobes may wish to skip, we will show how double left adjoints can help to analyze clitic (i.e., unstressed) pronouns in French. To start with, consider the clitic pronoun

$$\textit{le} : \bar{\mathbf{i}}\mathbf{o}^{\ell\ell}\bar{\mathbf{i}}^{\ell}.$$

Here we have introduced a new basic type $\bar{\mathbf{i}}$, strictly larger than $\mathbf{i}$ and postulate

$$\bar{\mathbf{i}} \not\to \mathbf{i} \to \bar{\mathbf{i}}.$$

The purpose of the overbar will be made clear later.

We begin with some sample sentences:

$$\begin{array}{c}\textit{je veux dormir}\\ \pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^{\ell})\ \mathbf{i} \quad \to \mathbf{s}_1\\ \textit{je veux voir Jean}\\ \pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^{\ell})\ (\mathbf{io})^{\ell})\ \mathbf{o} \quad \to \mathbf{s}_1\end{array}$$

The name *Jean* here appears as a direct object, but it could also be the subject of a sentence. Had we assigned the type $\mathbf{n}$ to names, we should have postulated

$$\mathbf{n} \to \mathbf{o}, \qquad \mathbf{n} \to \pi_3.$$

We are now able to handle

$$\begin{array}{c}\textit{je veux le voir —}\\ \pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^{\ell})\ (\bar{\mathbf{i}}\mathbf{o}^{\ell\ell}\bar{\mathbf{i}}^{\ell})\ (\mathbf{io}^{\ell}) \quad \to \mathbf{s}_1.\end{array}$$

To explain the purpose of the bar, we introduce another clitic pronoun

$$\textit{lui} : \bar{\mathbf{i}}\omega^{\ell\ell}\mathbf{i}^{\ell}$$

where

$$\omega = \text{indirect object}$$

is a new basic type. We assign two types to the verb *donner*, namely

$$\textit{donner} : \mathbf{i}\omega^{\ell}\mathbf{o}^{\ell}, \qquad \mathbf{io}^{\ell}\omega^{\ell}$$

to justify the sentences

$$\begin{array}{c}\textit{je veux donner (un livre) (à Jean)}\\ \pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^{\ell})\ (\mathbf{i}\omega^{\ell}\mathbf{o}^{\ell})\quad \mathbf{o} \quad\quad \omega \quad \to \mathbf{s}_1.\end{array}$$

and

$$\begin{array}{c}\textit{je veux donner (à Jean) (un livre)}\\ \pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^{\ell})\ (\mathbf{io}^{\ell}\omega^{\ell})\quad \omega \quad\quad \mathbf{o} \quad \to \mathbf{s}_1.\end{array}$$

[5]The reader may wonder why I did not take $\hat{\mathbf{o}} = \mathbf{o}$, but the explanation would take us too far afield.

Now consider

je veux le donner (à Jean)

$$\pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^\ell)\ (\bar{\mathbf{i}}\mathbf{o}^{\ell\ell}\bar{\mathbf{i}}^\ell)\ (\mathbf{i}\mathbf{o}^\ell\omega^\ell)\quad \omega \quad \to \mathbf{s}_1$$

je veux lui donner (un livre)

$$\pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^\ell)\ (\bar{\mathbf{i}}\omega^{\ell\ell}\mathbf{i}^\ell)\ (\mathbf{i}\omega^\ell\mathbf{o}^\ell)\quad \mathbf{o} \quad \to \mathbf{s}_1$$

je veux le lui donner

$$\pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^\ell)\ (\bar{\mathbf{i}}\mathbf{o}^{\ell\ell}\bar{\mathbf{i}}^\ell)\ (\bar{\mathbf{i}}\omega^{\ell\ell}\mathbf{i}^\ell)\ (\mathbf{i}\omega^\ell\mathbf{o}^\ell) \to \mathbf{s}_1$$

but

* *je veux lui le donner*

$$\pi_1\ (\pi_1^r\mathbf{s}_1\bar{\mathbf{i}}^\ell)\ (\bar{\mathbf{i}}\omega^{\ell\ell}\mathbf{i}^\ell)\ (\bar{\mathbf{i}}\mathbf{o}^{\ell\ell}\bar{\mathbf{i}}^\ell)\ (\mathbf{i}\mathbf{o}^\ell\omega^\ell) \not\to \mathbf{s}_1$$

Note that linguists usually put an asterisk on the left of an incorrect sentence. For a fuller treatment of French, the interested reader may wish to consult [1].

German à la Mark Twain

Not wishing to assume that readers of this article are familiar with German, I will avail myself of a trick due to Mark Twain [29]. He employed it to illustrate the vagaries of German by using English words with German word order. It is my contention that this strange word order will be triggered by the types assigned to the verbs. Here are some examples:

you see him

$$\pi_2\ (\pi_2^r\mathbf{s}_1\mathbf{o}^\ell)\ \mathbf{o} \quad \to \mathbf{s}_1$$

see you him?

$$(\mathbf{q}_1\mathbf{o}^\ell\pi_2^\ell)\ \pi_2\ \mathbf{o} \quad \to \mathbf{q}_1$$

The metarule for forming questions, which applies to English auxiliary verbs, applies to all verbs in German.

I have him seen

$$\pi_1\ (\pi_1^r\mathbf{s}_1\mathbf{p}_2^\ell)\ \mathbf{o}\ (\mathbf{o}^r\mathbf{p}_2) \quad \to \mathbf{s}_1$$

I can him see

$$\pi_1\ (\pi_1^r\mathbf{s}_1\mathbf{i}^\ell)\ \mathbf{o}\ (\mathbf{o}^r\mathbf{i}) \quad \to \mathbf{s}_1$$

he can seen become

$$\pi_3\ (\pi_3^r\mathbf{s}_1\mathbf{i}^\ell)\ (\mathbf{o}^r\mathbf{p}_2)\ (\mathbf{p}_2^r\mathbf{o}^{rr}\mathbf{i}) \quad \to \mathbf{s}_1$$

can he seen become ?

$$(\mathbf{q}_1\mathbf{i}^\ell\pi_3^\ell)\ \pi_3\ (\mathbf{o}^r\mathbf{p}_2)\ (\mathbf{p}_2^r\mathbf{o}^{rr}\mathbf{i}) \quad \to \mathbf{q}_1$$

Here the English *become* is used to translate the German passive auxiliary. The above analysis follows [16] in its use of double right adjoints. However, a different approach [22] shows that double right adjoints can be avoided in this context, by assigning a second type to the past participles of transitive verbs. Say, with the help of an appropriate metarule we obtain

$$\textit{seen} : \mathbf{p}_2\hat{\mathbf{o}}^\ell,$$

where $\hat{\mathbf{o}}$ is a new basic type, strictly smaller than $\mathbf{o}$, that is, such that

$$\mathbf{o} \not\to \hat{\mathbf{o}} \to \mathbf{o}.$$

The hat here guards against the following, which is correct in English, but not in German:

* *I have seen him*

$$\pi_1\ (\pi_1^r\mathbf{s}_1\mathbf{p}_2^\ell)\ (\mathbf{p}_2\hat{\mathbf{o}}^\ell)\ \mathbf{o} \quad \not\to \mathbf{s}_1.$$

Note that $\hat{\mathbf{o}}^\ell\mathbf{o} \to 1$ would imply

$$\mathbf{o} = 1\mathbf{o} \to \hat{\mathbf{o}}\hat{\mathbf{o}}^\ell\mathbf{o} \to \hat{\mathbf{o}}1 = \hat{\mathbf{o}}.$$

By assigning the following revised type to the passive auxillary

$$\textit{become} : \hat{\mathbf{o}}^r\mathbf{p}_2^r\mathbf{i},$$

we can re-analyze

he can seen become

$$\pi_3\ (\pi_3^r\mathbf{s}_1\mathbf{i}^\ell)\ (\mathbf{p}_2\hat{\mathbf{o}}^\ell)\ (\hat{\mathbf{o}}\mathbf{p}_2^r\mathbf{i}) \quad \to \mathbf{s}_1$$

without using double right adjoints.

Some Limitations to Our Approach

Other languages for which a preliminary pregroup analysis has been applied are Italian [7], Polish [12], Japanese [6] and Arabic [2]. Work on Latin and Turkish is in progress. However, evidence for double adjoints has so far been uncovered only in modern European languages.

We have been working with the free pregroup generated by partially ordered sets. Thus we admit postulates of the form $\alpha \to \beta$ only when α and β are basic types. As long as this restriction is borne in mind, the Lemma offers a decision procedure for sentence verification. However, it is doubtful whether all linguistic phenomena can be handled successfully using free pregroups only. As Buszkowski [4] formally proved, grammars based on free pregroups are context-free. It is known [26] that certain languages, most familiarly Dutch, are not context-free. The usual proof of this relies on the well-known argument in formal language theory that the intersection of two context-free languages need not be context-free. This suggests that one should incorporate the intersection symbol into the pregroup, that is, to work with free lattice pregroups. As far as I know, such a project has not yet been attempted.[6]

Even in English, there are some problems with our approach. Consider the noun phrase

people whom I know —

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

Here I have used the types

- $\mathbf{p}$ = plural noun phrase,
- $\mathbf{s}$ = declarative sentence when the tense is irrelevant and the postulate

$$\mathbf{s}_k \to \mathbf{s}.$$

The relative pronoun, in this context

$$\textit{whom} : \mathbf{p}^r\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell,$$

[6] *Lattice pregroups* are pregroups endowed with a binary operation *meet* ($\wedge$) such that $x \to a \wedge b$ if and only if $x \to a$ and $x \to b$.

will have to be assigned different types when the relative pronoun modifies singular nouns or when it occurs in a non-restrictive relative clause. A problem arises when the word *whom* is omitted, as in the perfectly acceptable noun phrase

people ∅ I know.

One does not like to attach the type $\mathbf{p}^r\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$ to the empty string ∅, nor does one want to accept a solution involving pregroups which are not freely generated, e.g., by postulating

$$\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell \to \mathbf{p}.$$

The solution I now tend to favour is to adopt a metarule which states that any plural noun of type $\mathbf{p}$ also has type $\mathbf{p}\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$. Putting this in another way, we might say that all nouns (not just plural ones) have optional *invisible* endings of type $\hat{\mathbf{o}}^{\ell\ell}\mathbf{s}^\ell$.

Omitting relative pronouns may lead to sentences which are hard to analyze. For example, consider the two sentences

police whom police control control police,
police control police whom police control.

Omitting the relative pronoun *whom* and replacing the verb *control* by the synonymous verb *police*, we transform both sentences into

police police police police police.

Let me end this section with a problem for readers who like puzzles: Show that the string $police^{2n+3}$ (where $n \in \mathbb{N}$) can be parsed as a grammatical sentence in precisely $n!$ distinct ways.

Some Mathematical Examples

When I introduced pregroups for linguistic purposes at a conference in 1998, I did not immediately realize that they were not quite as new as I thought. In fact, I had exploited the pregroup of unbounded order-preserving mappings $\mathbb{Z} \to \mathbb{Z}$ myself in [14]. In the following discussion, let me talk about *left pregroups*, that is, partially ordered monoids in which each element has a left adjoint, but not necessarily a right adjoint. A left pregroup is easily seen to be a partially ordered group if and only if $a^{\ell\ell} = a$ for each element a. An example of a left pregroup which is not a group is the monoid of order-preserving unbounded mappings $\mathbb{N} \to \mathbb{N}$. In fact, this example occurs implicitly in a paper [21] written in collaboration with my late friend Leo Moser, although we were then innocent of the present terminology and its categorical connection. We did, however, study the following examples, among others, which I present here with streamlined arguments in modern terminology. (Similar results hold for pregroups, see [14] and [18].)

Consider the monoid of unbounded order-preserving functions $f: \mathbb{N} \to \mathbb{N}$, the binary operation being composition:

$$(gf)(x) = g(f(x)).$$

Define the partial order elementwise:

$$f \to g \text{ iff } \forall_{x \in \mathbb{N}} f(x) \le g(x),$$

and construct the left adjoint thus:

$$(*) \qquad f^\ell(x) \le y \text{ iff } x \le f(y),$$

that is,

$$f^\ell(x) = \inf\{y \in \mathbb{N} \mid x \le f(y)\}.$$

Some "universal" algebraists will recognize the pair (f^ℓ, f) as a *Galois connection* (see, e.g., [14] and [18]). It is easily seen that f^ℓ is indeed the left adjoint of f in the sense defined earlier. Indeed, putting $x = f(y)$, we see that $(f^\ell f)(y) \le y$ and, putting $y = f^\ell(x)$, we see that $x \le ff^\ell(x)$. To show that this left pregroup is not a group, take $f(x) = [x/2]$, then $f^\ell(x) = 2x$ and

$$f^{\ell\ell}(x) = [(x+1)/2],$$

which is $\neq f(x)$ in general.

Example 1: A left adjoint of interest in number theory is offered by the function

$$\pi(x) = \text{the number of primes} \le x.$$

Then $\pi^\ell(0) = 0$ and, for $x \ge 1$,

$$\pi^\ell(x) = p(x) = \text{the } x\text{th prime}.$$

Put $p(0) = \pi^\ell(0) = 0$ and look at Table 1.

Table 1

x	*p*(*x*)	*p*(*x*) + *x*	π(*x*)	π(*x*) + *x* + 1
0	0	0	0	1
1	2	3	0	2
2	3	5	1	4
3	5	8	2	6
4	7	11	2	7
5	11	16	3	9
...	...	...	...	...

Inspection of this table leads to the curious observation first made in [21]: the sets

$$\{p(x) + x \mid x \in \mathbb{N}\}, \quad \{\pi(x) + x + 1 \mid x \in \mathbb{N}\}$$

are complementary subsets of $\mathbb{N}$. Well, the numbers 10 and 12 to 15 seem to be missing, so they should appear in the last column further down, as the reader will easily verify.

A deep property of prime numbers? Not at all. As was shown in [21], we have the following general result:[7]

[7] In [21], we had not realized the importance of adjoints or Galois connections, so we considered f^ℓ to be a kind of "inverse" of f. Moreover, we removed the restriction of unboundedness by adjoining an infinite element to $\mathbb{N}$. A very readable account of this result and its applications will be found in [11].

PROPOSITION *For any unbounded order-preserving function $f: \mathbb{N} \to \mathbb{N}$, the sets*

$$(^{**}) \qquad \{f^{\ell}(x) + x \mid x \in \mathbb{N}\}, \quad \{f(x) + x + 1 \mid x \in \mathbb{N}\}$$

are complementary: every natural number belongs to one and only one of the two sets.

Since I won't expect the reader to look up an old issue of the *American Mathematical Monthly*, here is a recapitulation of the proof, which is quite easy, though a little tricky.

PROOF. It follows from (*) that

$$f^{\ell}(x) + x \leq xy \text{ iff } x + y \leq f(y) + y,$$

hence $f^{\ell}(x) + x \neq f(y) + y + 1$, and so the two sets under consideration have no elements in common. □

Now consider natural numbers x and y such that $x + y = n$, that is,

$$0 \leq x \leq n, \; y = n - x.$$

Then

$$f^{\ell}(x) + x \leq n \text{ iff } f(y) + y + 1 > n,$$

hence the two sets

$$\{f^{\ell}(x) + x \mid 0 \leq x \leq n\}, \quad \{f(y) + y + 1 \mid 0 \leq y \leq n\}$$

together have exactly $n + 1$ elements between 0 and n.

Since n is arbitrary, it follows that the two subsets (**) of $\mathbb{N}$ are complementary.

EXAMPLE 2: Here is another illustration of the Proposition. Let $f(x) = \{\sqrt{x+1}\}$ be the closest integer to the square-root of $x + 1$. Then $f^{\ell}(x) = x^2 - x$, and so $x + 1 + \{\sqrt{x+1}\}$ enumerates the natural numbers which are not perfect squares.

EXAMPLE 3: An even earlier example preceded the Proposition and had, in fact, inspired it. Let $f(x) = [\rho x]$ be the greatest integer $\leq \rho x$, ρ being an irrational number. Then $f^{\ell}(x) = [\rho^{-1} x]$ and so

$$\{[(\rho^{-1} + 1)x] \mid x \in \mathbb{N}\}, \quad \{[(\rho + 1)(x + 1)] \mid x \in \mathbb{N}\}$$

are complementary subsets of $\mathbb{N}$. Discarding $x = 0$ in the first of these sets and replacing $x + 1$ by x in the second, we find that

$$\{[(\rho^{-1} + 1)x] \mid x \geq 1\}, \quad \{[(\rho + 1)x] \mid x \geq 1\}$$

are complementary sets of positive integers. This ancient result by Beatty first appeared as a problem in the *American Mathematical Monthly* [3].

EXAMPLE 4: A special case of Beatty's problem involves the golden ratio τ satisfying $\tau^2 = \tau + 1$, for which

$$\{[\tau^2 x] \mid x \geq 1\}, \quad \{[\tau x] \mid x \geq 1\}$$

are complementary sets of positive integers. These play a rôle in describing the winning strategy of Wythoff's game [25].

Related Mathematical and Logical Systems

The ordered algebraic system studied here is that of a pregroup. Many years ago [13], I had suggested another system,[8] that of a *residuated monoid*, that is, a partially ordered monoid with two binary operations/(over) and\(under) satisfying

$$ab \to c \text{ iff } a \to c/b \text{ iff } b \to a\backslash c.$$

A residuated monoid with lattice operations becomes a *Grishin algebra* if we add a so-called dualizing element 0 such that

$$(0/a)\backslash 0 = a = 0/(a\backslash 0).$$

Then one can define a De Morgan dual + to the operation here denoted by juxtaposition:

$$a + b = ((0/b)(0/a))\backslash 0$$

or, equivalently,

$$a + b = 0/((b\backslash 0)\,(a\backslash 0)).$$

A pregroup may then be viewed as a Grishin algebra in which

$$0 = 1, \; a + b = ab,$$

as is easily seen by defining

$$a^{\ell} = 0/a, \quad a^{r} = a\backslash 0.$$

On the other hand, a residuated monoid may also be turned into a semi-Heyting algebra, by introducing the postulates[9]

$$a \to 1, \quad a \to aa, \quad ab \to ba.$$

For logicians, free residuated monoids, Grishin algebras, pregroups, and semi-Heyting algebras correspond to the following logical systems respectively:

syntactic calculus,
classical bilinear logic,
compact bilinear logic,
positive intuitionistic propositional calculus.

I have been asked what led me to pregroup grammars and why I prefer them to the earlier syntactic calculus [13] and its recent offshoots [24]. It was Claudia Casadio who had the pioneering idea of introducing non-commutative linear logic (classical bilinear logic) for linguistic purposes (see the discussion in [8]). This system employed, in addition to a tensor product, also its De Morgan dual. Once one realizes that there is no need for two binary operations in grammatical applications, one is led to compact bilinear logic, which identifies the two.

[8]Originally I had considered a *residuated semigroup*, but later decided to add a unity element.

[9]These are the algebraic versions of Gerhard Gentzen's three structural rules: *weakening, contraction* and *interchange*. In this algebraic version the last may be derived from the other two: $ab \to abab \to ba$.

The original syntactic calculus and its more recent descendants display grammatical derivations as page-filling proof trees. On the other hand, computations in pregroup grammars are one-dimensional and can be carried out in real time, without overburdening the short-term memory, hence my present preference for the latter system.

REFERENCES

1. D. Bargelli and J. Lambek, An algebraic approach to French sentence structure, in: P. de Groote *et al.* (eds), *Logical aspects of computational linguistics*, Springer LNAI 2099 (2001), 62–78.
2. D. Bargelli and J. Lambek, An algebraic approach to Arabic sentence structure, *Linguistic Analysis* 31 (2001), 301–315.
3. S. Beatty, Problem 3173, *Amer. Math. Monthly* 33 (1926), 159.
4. W. Buszkowski, Lambek grammars based on pregroups, in: P. de Groote *et al.* (eds), *Logical aspects of computational linguistics*, Springer LNAI 2099 (2001), 95–109.
5. W. Buszkowski, Cut elimination for the Lambek calculation with adjoints, in: V. M. Abrusci *et al.* (eds), *New perspectives in logic and formal linguistics*, Bulzoni Editore, Rome, 2002, 85–93.
6. K. Cardinal, *An algebraic study of Japanese grammar*, MSc. Thesis, McGill University, 2002.
7. C. Casadio and J. Lambek, An algebraic analysis of clitic pronouns in Italian, in: P. de Groote *et al.* (see [1]), 116–124.
8. C. Casadio and J. Lambek, A tale of four grammars, *Studia Logica* 71 (2002), 315–329.
9. A. Glass, *Partially ordered groups*, World Scientific, Singapore, 1998.
10. Z. Harris, A cyclic cancellation-automaton for sentence well-formedness, *International Computation Centre Bulletin* 5 (1966), 69–94.
11. R. Honsberger, *Ingenuity in mathematics*, New Mathematical Library 23, Mathematical Association of America 1970, Essay 12.
12. A. Kiślak, Pregroups versus English and Polish grammar, in: P. de Groote *et al.* (see [1]), 129–154.
13. J. Lambek, The mathematics of sentence structure, *Amer. Math. Monthly* 65 (1958), 154–170.
14. ——, Some Galois connections in elementary number theory, *J. Number Theory* 47 (1994), 371–377.
15. ——, Type grammar revisited, in: A. Lecomte *et al.* (eds), *Logical Aspects of Computational Linguistics*, Springer LNAI 1582 (1999), 1–27.
16. ——, Type grammar meets German word order, *Theoretical Linguistics* 26 (2000), 19–30.
17. ——, Pregroups: a new algebraic approach to sentence structure, in: V. M. Abrusci *et al.* (see [5]), 39–54.
18. ——, Iterated Galois connections in arithmetic and linguistics, in: K. Denecke *et al.* (eds), *Galois connections and applications*, Mathematics and its Applications, Volume 565, Kluwer, 2004.
19. ——, What is a pregroup? in: C. Casadio, P. J. Scott and R. A. G. Seely (eds), Laguage and grammar: *Studies in mathematical linguistics and natural language*, CSLI Press, Stanford University, 2005.
20. ——, A computational algebraic approach to English grammar, *Syntax* 7:2 (2004), 128–147.
21. J. Lambek and L. Moser, Inverse and complementary sequences of natural numbers, *Amer. Math. Monthly* 61 (1954), 454–458.
22. J. Lambek and A. Preller, An algebraic approach to the German sentence, *Linguistic Analysis* 31 (2001).
23. 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.
24. M. Moortgat, Categorical type logics, in: J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language*, Elsevier, Amsterdam, 1997, 93–177.
25. W. W. Rouse Ball and H. S. M. Coxeter, *Mathematical recreations and essays*, London, 1944.
26. S. Sheiber, Evidence against the context-freeness of natural languages, *Linguistics and Philosophy* 1985, 333–343.
27. C. S. Peirce, The logic of relatives, *The Monist* 71 (1887), 161–217.
28. A. S. Troelstra, *Lectures in Linear Logic*, CSLI Lecture Notes, Stanford, 1992.
29. M. Twain, The horrors of the German Language, in: *Mark Twain's speeches*, Harper and Row, New York, 1910, 45–51.