Joachim Lambek

Should Pregroup Grammars be Adorned with Additional Operations?

To Michael Moortgat on his first half century.

Even as I am trying to explain the pregroup approach to mainstream linguists [17], I note that there is still some concern from within the categorial grammar community as to whether the algebraic machinery of pregroups suffices for linguistic applications. Some of my friends would prefer to return to residuated monoids or even residuated semigroups [12], thus preferring two binary operations of division, called "residuation", to the two unary operations, called "adjoints", which generalize the notion of inverse in a group. They may even question the associative law, as I did at one time [13], and suggest that associativity, when present, should be licensed by special operations, called "modalities" [22]. More recently [24], it has been suggested that the pregroup approach too may profit if such operations are adjoined.

I would like to take this opportunity to present my reasons for resisting the introduction of additional operations. Although I am primarily a mathematician and believe that grammar has more right to be called a branch of mathematics than the Pythagorean music or astronomy, I feel that the mathematics underlying the structure of natural language should be kept as simple as possible. In fact, even the pregroup approach appears to be too sophisticated for many, if not most languages. In Appendix I, I will present a possible scenario as to how the pregroup machinery could have risen in relatively recent times in Western Europe.

For readers not familiar with the notion of a pregroup[1], here is a brief résumé: a *pregroup* is a partially ordered monoid with two additional unary operations $(-)^l$ and $(-)^r$, called *left adjoint* and *right adjoint* respectively,

[1] I first introduced pregroups, though not under this name and without envisaging a linguistic application in [14].

Special Issue Categorial Grammars and Pregroups
Edited by **Wojciech Buszkowski and Anne Preller**

Studia Logica (2007) 87: 343–358
DOI: 10.1007/s11225-007-9094-1

subject only to the rules

$$a^l a \to 1 \to a a^l,\ a a^r \to 1 \to a^r a,$$

where the arrow denotes the partial order.

The main idea is to work with the *free* pregroup generated by a partially ordered set of *basic types*. Its elements, just called *types*, are strings of *simple types* of the form

$$\cdots, a^{ll}, a^l, a, a^r, a^{rr}, \cdots,$$

a being a basic type. One is supposed to assign to each word in the (mental) dictionary one or more types and use the algebraic machinery for calculating the type of a string of words. As long as one wishes to check whether the resulting type is a simple type, it suffices to use the *contractions* $a^l a \to 1$ and $a a^r \to 1$. This was proved in [16] using what other people (e.g. [4]) have called the "switching lemma", which allows contractions to precede expansions $1 \to a a^l$ and $1 \to a^r a$.

In their recent paper [24] Moortgat and Oehrle say this: "The pregroup grammar framework in itself, however, is not expressive enough to adequately analyze natural language syntax and semantics". They may well be right about the syntax, although I have not been convinced by the examples they have chosen. They are certainly right about the semantics, but even here there are attenuating circumstances, which I will discuss in Appendix II.

They point out correctly that, if adjectives are assigned the type $\mathbf{n}\mathbf{n}^l$, where

$$\mathbf{n} = \text{(type of) noun},$$

and if the adverb *very* is assigned the type

$$(\mathbf{n}\mathbf{n}^l)(\mathbf{n}\mathbf{n}^l)^l = \mathbf{n}\mathbf{n}^l\mathbf{n}^{ll}\mathbf{n}^l,$$

the latter type may be contracted to $\mathbf{n}\mathbf{n}^l$, wrongly allowing the adverb *very* to be treated as an adjective.

As I have elaborated elsewhere [18], the English adjective, like the adjective in many languages, serves two distinct functions: it can occur in a predicate, as in *"he is smart"*, or it can occur as an attribute, as in *"the smart man"*.

Let us adopt the basic type

$$\mathbf{a} = \text{(type of) adjective}$$

and assign to the adverb *very* the type $\mathbf{a}^l\mathbf{a}$, thus allowing the calculation

$$\begin{array}{c} \textit{he is very smart} \\ \pi_3(\pi_3^r\mathbf{s}_1\mathbf{a}^l)(\mathbf{a}\mathbf{a}^l)\mathbf{a} \to \mathbf{s}_1 \end{array}$$

where

$$\begin{array}{rl} \pi_3 = & \text{third person singular subject,} \\ \mathbf{s}_1 = & \text{declarative sentence in present tense} \end{array}$$

and the copula *is* has been assigned the type $\pi_3^r\mathbf{s}_1\mathbf{a}^l$. Now how do we handle *"the very smart man"?* As a first attempt, we might adopt the provisional metarule:

every adjective of type $\mathbf{a}$ also has type $\mathbf{n}^l\mathbf{n}$.

This would explain *"the smart man"*, but not *"the very smart man".*

Note that this problem does not arise in German, a language closely related to English. In German, there is a morphological distinction between the adjective used predicatively and the adjective used attributively:

$$\begin{array}{cc} \textit{er ist klug} & \textit{der klug + e Mann} \\ \pi_3(\pi_3^r\mathbf{s}_1\mathbf{a}^l)\mathbf{a} \to \mathbf{s}_1 & (\pi_3\mathbf{n}^r)\mathbf{a}(\mathbf{a}^r\mathbf{n}\mathbf{n}^l)\mathbf{n} \to \pi_3 \end{array}$$

where the morpheme *+e* has been assigned the type $\mathbf{a}^r\mathbf{n}\mathbf{n}^l$. (The German adjective also admits other endings *+em, +en, +er, +es* to encode other information; see [19] for details.) There is no problem if we introduce the adverb *sehr* of type $\mathbf{a}\mathbf{a}^l$ before *klug* in either case.

In English, we can adopt the same strategy, by saying that every adjective can have an *invisible* ending of type $\mathbf{a}^r\mathbf{n}\mathbf{n}^l$. Equivalently, we might adopt the revised metarule:

every adjective of type $\mathbf{a}$ also has type $\mathbf{a}(\mathbf{a}^r\mathbf{n}\mathbf{n}^l)$.

Since $\mathbf{a}\mathbf{a}^r\mathbf{n}\mathbf{n}^l \to \mathbf{n}\mathbf{n}^l$, this implies that our provisional metarule is still valid. But now the revised metarule allows the following calculations:

$$\begin{array}{cc} \textit{he is very smart} & \textit{the very smart man} \\ \pi_3(\pi_3^r\mathbf{s}_1\mathbf{a}^l)(\mathbf{a}\mathbf{a}^l)\mathbf{a} \to \mathbf{s}_1 & (\bar{\mathbf{n}}\mathbf{n}^l)(\mathbf{a}\mathbf{a}^l)\mathbf{a}(\mathbf{a}^r\mathbf{n}\mathbf{n}^l)\mathbf{n} \to \bar{\mathbf{n}} \end{array}$$

where

$$\bar{\mathbf{n}} = \text{complete singular noun phrase}$$

and $\bar{\mathbf{n}} \to \pi_3$, recalling that the basic types form a partially ordered set.

In principle, such metarules are not required, since the dictionary could list both types $\mathbf{a}$ and $\mathbf{a}(\mathbf{a}^r\mathbf{n}\mathbf{n}^l)$ for each adjective. In practice, they help to prevent overloading the dictionary.

Moortgat and Oehrle [24] discuss a second example, quoted from [13], to justify departure from associativity:

** who works and John rests?*

It had been noted that, already in the associative syntactic calculus [12], what seemed to be the most natural assignment would falsely accept this as a grammatical sentence. The same is true if we assign to *who* the type $\mathbf{q}\mathbf{s}^l\pi_3$ in a pregroup grammar, where

$$\begin{aligned} \mathbf{q} &= \quad \text{(type of) question,} \\ \mathbf{s} &= \quad \text{declarative sentence} \end{aligned}$$

and $\mathbf{s}_1 \to \mathbf{s}$. This would accept correctly

$$\begin{gathered} \textit{who works} \\ (\mathbf{q}\mathbf{s}^l\pi_3)(\pi_3^r\mathbf{s}_1) \to \mathbf{q} \end{gathered}$$

but incorrectly

$$\begin{gathered} \textit{* who works and John rests?} \\ (\mathbf{q}\mathbf{s}^l\pi_3)(\pi_3^r\mathbf{s}_1)(\mathbf{s}^r\mathbf{s}\mathbf{s}^l)\pi_3(\pi_3^r\mathbf{s}_1) \to \mathbf{s} \end{gathered}$$

Well, I believe that *who* should be given a different type altogether, namely $\mathbf{q}\hat{\pi}_3^{ll}\mathbf{q}^l$. I would not derive *"who works?"* form *"he works"*, but from the pseudo-question *"*works he?"* In fact, Chomsky himself appears to have done something like this in [7] [section 7.2 (60)]. Again, in German the latter would actually be acceptable.

I would now assign to *works* a second type $\mathbf{q}_1\hat{\pi}_3^l$ (with the help of an appropriate metarule), where

$$\begin{aligned} \mathbf{q}_1 &= \quad \text{question in the present tense,} \\ \hat{\pi}_3 &= \quad \text{pseudo-subject} \end{aligned}$$

and postulate $\mathbf{q}_1 \to \mathbf{q}$, but $\pi_3 \not\to \hat{\pi}_3$. This will account for the difference in acceptability between the following:

$$\begin{array}{cc} \textit{who works - ?} & \textit{* works he ?} \\ (\mathbf{q}\hat{\pi}_3^{ll}\mathbf{q}^l)(\mathbf{q}_1\hat{\pi}_3^l) \to \mathbf{q} & (\mathbf{q}_1\hat{\pi}_3^l)\pi_3 \not\to \mathbf{q}_1 \end{array}$$

The former, with a Chomskyan trace at the end, is justified because $\mathbf{q}_1 \to \mathbf{q}$. To justify the latter we would require $\hat{\pi}_3^l \pi_3 \to 1$, but then $\pi_3 \to \hat{\pi}_3 \hat{\pi}_3^l \pi_3 \to \hat{\pi}_3$, contrary to the assumption that $\pi_3 \to \hat{\pi}_3$ has not been postulated. (This is what I mean by writing $\pi_3 \not\to \hat{\pi}_3$.)

Now return to

$$\begin{array}{c} *\ \textit{who works and John rests ?} \\ (\mathbf{q}\hat{\pi}_3^{ll}\mathbf{q}^l)(\mathbf{q}_1\hat{\pi}_3^l)(x^r x x^l)\pi_3(\pi_3^r \mathbf{s}_1) \to \mathbf{s}_1 \end{array}$$

There is no choice of x yielding an acceptable type assignment to the conjunction *and* which would allow the co-ordination of a question with a declarative sentence.

Moortgat and Oehrle have developed an ingenious strategy to overcome the shortcomings of traditional categorial grammars by endowing their non-associative systems with powerful modalities, which will license associativity or commutativity when needed. Of course, pregroups are associative to start with, but how are we to justify occasional commutativity? In my opinion, this can be done most easily by introducing multiple type assignment into the dictionary or, equivalently, with the help of appropriate metarules. For example, all six permutations of the Latin sentence

$$\begin{array}{ccc} \textit{puer} & \textit{puellam} & \textit{amat} \\ \pi_3 & \mathbf{o} & (\mathbf{o}^r \pi_3^r \mathbf{s}_1) \end{array}$$

where

$$\mathbf{o} = \text{(type of) object},$$

are permitted in principle, although with different emphases. We would assign to *amat* the principal type $\mathbf{o}^r \pi_3^r \mathbf{s}_1$ and allow appropriate metarules to produce the additional types

$$\pi_3^r \mathbf{o}^r \mathbf{s}_1, \quad \mathbf{s}_1 \pi_3^l \mathbf{o}^l, \quad \pi_3^r \mathbf{s}_1 \mathbf{o}^l, \quad \mathbf{s}_1 \mathbf{o}^l \pi_3^l, \quad \mathbf{o}^r \mathbf{s}_1 \pi_3^l.$$

(See [6] for more details).

A word of caution: actually a Latin dictionary does not list the form *amat*, but the first person *amo*, hopefully with sufficient information to calculate the 90 finite inflected forms of the verb. Even an English dictionary does not list *works*, but the infinite *work*, from which the finite forms *works* and *worked* may be derived.

Another example considered by Moortgat and Oehrle to illustrate their modalities is the following, which I prefer to analyze without modalities:

$$\textit{what Alice found there -}$$
$$(@) \quad (\bar{\mathbf{n}}\hat{\mathbf{o}}^{ll}\mathbf{s}^l)\bar{\mathbf{n}}(\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l\lambda^l)\lambda \to \bar{\mathbf{n}}$$

where

$\hat{\mathbf{o}} =$	pseudo-object,
$\lambda =$	adverb (or prepositional phrase) of location,
$\bar{\mathbf{n}} =$	complete singular noun phrase,
$\mathbf{s}_2 =$	declarative sentence in the past tense,

subject to $\bar{\mathbf{n}} \to \pi_3$, $\bar{\mathbf{n}} \to \mathbf{o}$ and $\mathbf{s}_2 \to \mathbf{s}$.
The principal type of *found* is $\pi_3^r\mathbf{s}_2\mathbf{o}^l$, to justify

$$\textit{she found it / a bottle}$$
$$\pi_3(\pi_3^r\mathbf{s}_2\mathbf{o}^l)\mathbf{o} \to \mathbf{s}_2$$

Of course, the dictionary lists the type of the infinitive *find*, as well as the irregular past tense *found*. To justify

$$\textit{she found it there}$$
$$\pi_3(\pi_3^r\mathbf{s}_2\lambda^l\mathbf{o}^l)\mathbf{o}\lambda \to \mathbf{s}_2$$

we must cite an appropriate metarule to derive the new type $\pi_3^r\mathbf{s}_2\lambda^l\mathbf{o}^l$ for *found*. Finally, to justify the noun phrase (@), we first introduce the pseudo-sentence

$$\textit{* she found there it}$$
$$\pi_3(\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l\lambda^l)\lambda\mathbf{o} \not\to \mathbf{s}_2$$

where *found* has been assigned yet another type and $\mathbf{o} \not\to \hat{\mathbf{o}}$. We can now analyze

$$\textit{what she found there}$$
$$(\bar{\mathbf{n}}\hat{\mathbf{o}}^{ll}\mathbf{s}^l)\pi_3(\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l\lambda^l)\lambda \to \bar{\mathbf{n}}$$

by assigning the appropriate type to *what*.

To compare the pregroup approach with the multimodal one, Moortgat and Oehrle say in their [24] conclusion: “to move beyond context free expressitivity, the two grammar formalisms each have their own strategy: closing the lexicon under metarules in the case of pregroup grammars; adding lexically controlled structural postulates in the case of type-logical grammars.” As far as I can see, the metarules for pregroup grammars construct only a finite number of additional types in the dictionary, so they do not allow us to move beyond “context free expressivity”. To do so, I would follow the lead of Kanazawa [10] and introduce lattice operations into the pregroup.

This is easily illustrated with the help of formal languages, although the pregroup approach was designed for natural, not formal languages. For example, consider the well-known context-free language

$$L = \{A^m B^n C^n | m, n \geq 1\}.$$

Its sentences can be recognized by a pregroup grammar with basic types

$$\mathbf{s} = \text{sentence}, \ \mathbf{t}, \ \mathbf{c}$$

and the following type assignments

$$\begin{array}{l} A : \mathbf{st}^l, \mathbf{st}^l\mathbf{t}^{ll}\mathbf{s}^l, \\ B : \mathbf{tc}^l, \mathbf{tc}^l\mathbf{t}^l, \\ C : \mathbf{c}. \end{array}$$

Let me declare: this is not the recommended way of describing L!

Now, a typical example of a non-context-free language is the intersection $L \cap L'$, where

$$L' = \{A^m B^m C^n | m, n \geq 1\}.$$

may be analyzed similarly to L, with sentences of type $\mathbf{s}'$. If we adopt the algebraic machinery of a lattice pregroup[2], i.e. a pregroup with an additional binary operation $\wedge$ such that

$$c \to a \wedge b \text{ if and only if } c \to a \text{ and } c \to b,$$

then we can easily describe the sentences of

$$L \cap L' = \{A^n B^n C^n | n \geq 1\}$$

as having type $\mathbf{s} \wedge \mathbf{s}'$.

This looks like a promising approach for handling natural languages such as Swiss German, which has been proved to be non-context-free [28]. Unfortunately, we are still waiting for a solution to the word problem for free lattice pregroups, analogous to the switching lemma for free pregroups. As Buszkowski [4], who has made a thorough investigation of this question, puts it: what is wanted is a "cut-free axiomatization for compact bilinear logic with lattice operations".

However, there is an alternative approach. As long as one is interested only in computation, there is no need to enlarge the algebra in order to

---

[2] I had called lattice pregroups "lattice ordered monoids with adjoints" in [15], again with no linguistic application in mind.

handle the intersection of two context-free languages. All one has to do is to perform two parallel calculations on the string of type symbols and observe that one contracts to $\mathbf{s}$, the other to $\mathbf{s}'$. But somehow I doubt that Swiss children proceed in this manner.

Mario Fadda [8] also wishes to introduce modalities into pregroup grammars to overcome problems with "some phenomena related with word order, discontinuous constituents, etc...". I do not know about the "etc", but I believe the other problems can be handled without modalities.

Compared to English, German has a strange word order, but this can be shown to be triggered by assigning appropriate types to words. For example, consider the following German sentence, with a paraphrase in the spirit of Mark Twain [29]:

wirst Du gesehen worden sein
= *will you seen got be*

$$(\mathbf{q}_1\mathbf{i}^l\pi_2^l)\pi_2(\mathbf{o}^r\mathbf{p}_2)(\mathbf{p}_2^r\mathbf{o}^{rr}\mathbf{p}_2')(\mathbf{p'}_2^r\mathbf{i}) \to \mathbf{q}_1$$

which may be evaluated by rebracketing thus:

$$\mathbf{q}_1[\mathbf{i}^l[\pi_2^l\pi_2][\mathbf{o}^r[\mathbf{p}_2\mathbf{p}_2^r]\mathbf{o}^{rr}][\mathbf{p'}_2\mathbf{p'}_2^r]\mathbf{i}].$$

Here we have introduced some new basic types:

$\mathbf{i}$	=	infinitive of intransitive verb phrase,
π_2	=	second person singular subject,
$\mathbf{p}_2$	=	past participle of verbs that go with *haben*,
$\mathbf{p'}_2$	=	past participle of verbs that go with *sein*.

For further details see [19], and [20] for an alternative approach avoiding double *right* adjoints.

How are discontinuous constituents handled in pregroup grammars? Consider the following examples:

he turned off the light,	** he turned off it,*
he turned the light off,	*he turned it off.*

In the first two examples, *turned off* is a constituent, though the second example is ungrammatical. In the last two sample sentences, there is a discontinuity between *turned* and *off*.

Evidently, the constituent *turned off* should have type $\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l$, where $\bar{\mathbf{n}} \to \hat{\mathbf{o}}$, but $\mathbf{o} \not\to \hat{\mathbf{o}}$:

he (turned off) (the light),	** he (turned off) it,*
$\pi_3(\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l)\bar{\mathbf{n}} \to \mathbf{s}_2$	$\pi_3(\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l)\mathbf{o} \not\to \mathbf{s}_2$

since $\hat{\mathbf{o}}^l\mathbf{o} \to 1$ would imply $\mathbf{o} \to \hat{\mathbf{o}}$. If we assign to *off* the type δ, *turned* will have type $\pi_3^r\mathbf{s}_2\hat{\mathbf{o}}^l\delta^l$.

To handle the last two sentences, *turned* requires a second type $\pi_3^r\mathbf{s}_2\delta^l\mathbf{o}^l$:

he turned (the light) off,	*he turned it off,*
$\pi_3(\pi_3^r\mathbf{s}_2\delta^l\mathbf{o}^l)\bar{\mathbf{n}}\delta \to \mathbf{s}_2$	$\pi_3(\pi_3^r\mathbf{s}_2\delta^l\mathbf{o}^l)\mathbf{o}\delta \to \mathbf{s}_2$

provided we postulate $\bar{\mathbf{n}} \to \mathbf{o}$.

There may indeed be good linguistic reasons for enhancing the syntactic power of pregroups with the help of additional operations. But, aside from the conceivable addition of lattice operations, I have seen no convincing evidence for this. This is not to say that the resulting elegant algebraic or logical systems may not themselves be worthy of study.

Appendix I.
A historical speculation.

Some of my categorial fellow travellers are enamoured of diamonds and boxes and would like to introduce these into pregroups. Mathematically, such additional unary operations, or the binary lattice operations discussed above, are certainly worthy of being studied. I have not yet been convinced though that they can help significantly to analyze sentences in natural languages that cannot already be analyzed more modestly in unadorned pregroup grammars. In fact, I wonder how even such a relatively sophisticated tool like a pregroup could have been adopted in certain linguistic communities. Let me propose a conjectural scenario how this might have come about.

A preliminary investigation has turned up no evidence for double adjoints in a number of languages, such as Arabic [2], Latin [6] or Turkish [3]. This is not to deny that such evidence may turn up when these languages are studied more thoroughly. In the absence of such evidence, the grammars of these languages can be analyzed with the help of a *rudimentary* technique that avoids abstract mathematics.

We begin with a partially ordered set of *basic types*. We assign to each word (or morpheme) a *type*, namely a string of *simple* types $\alpha_1 \ldots \alpha_n$, where each α_i has the form a, a^l or a^r, a being a basic type. When looking at a given string of words or morphemes, we also concatenate their types and perform repeated (generalized) contractions as follows, when $a \to b$:

$$b^l a \to 1, \quad ab^r \to 1,$$

where 1 is the empty string of simple types. Hopefully, the result of these repeated contractions will be a basic type telling us that the given string of words is a grammatical sentence.

This rudimentary kind of grammar seems to suffice for many languages. It is essentially what Zelig Harris proposed in [9].[3] It may even be implicit in the ideas of C.S. Peirce [25], who might have looked at the type $\pi_3^r \mathbf{s}_2 \lambda^l \mathbf{o}^l$ of *found* and compared π_3^r, λ^l and $\mathbf{o}^l$ to the *unsaturated bonds* of an atom.[4] I believe that pregroup grammars developed from this rudimentary setup.

Now I must confess that I cannot explain how or why a linguistic community will adopt a certain grammatical change, anymore than I can explain how or why a flock of pigeons will take off all at the same time. I do not wish to postulate a "collective subconscious", nor pretend that such decisions are made by a committee of grammarians, somewhat like the Academie Française. Not being trained in historical linguistics, I can only conjecture the following scenario.

At some time in history, perhaps in the early Renaissance, people in some linguistic communities realized that the partial order of basic types could be extended first to simple types by stipulating *the contravariance of adjunction*:

$$\text{if } a \to b \text{ then } b^l \to a^l \text{ and } b^r \to a^r,$$

and then to strings of simple types by demanding that

$$\text{if } \alpha \to \beta \text{ then } \ldots\ \alpha\ \ldots \to \ldots\ \beta\ \ldots\ .$$

These two requirements allow the generalized contractions to be derived from simple contractions

$$b^l b \to 1, \quad b b^r \to 1$$

by arguing, for example, as follows:

$$a b^r \to b b^r \to 1,$$

or even

$$a b^r \to a a^r \to 1.$$

Now it may have occurred to people that the operations $(-)^l$ and $(-)^r$ could be extended to arbitrary types, provided one admitted more simple types:

[3] I am indebted to A. Joshi for drawing my attention to this article.

[4] I thank Claudia Casadio for digging up the relevant article [25].

$$..., a^{ll}, a^{l}, a, a^{r}, a^{rr}, ...$$

and stipulated

$$1^l = 1 = 1^r, \quad (\alpha\beta)^l = \beta^l\alpha^l, \quad (\alpha\beta)^r = \beta^r\alpha^r.$$

Why the order reversal? Because now

$$(\alpha\beta)^l(\alpha\beta) = \beta^l\alpha^l\alpha\beta \to \beta^l\beta \to 1.$$

It also seemed reasonable to stipulate

$$\alpha^{lr} = \alpha = \alpha^{rl}$$

in order to prevent an explosion of the number of simple types.

While I have seen no evidence for a^{lll} or a^{rrr} in any language, a^{ll} (and perhaps a^{rr}) turned out to be useful in English, German, French, Italian and Polish for describing what Chomsky called "traces" and for analyzing clitic pronouns in Romance languages [1], [5].

Now what does this have to do with pregroups? We have already ensured that $\alpha^l\alpha \to 1$ and $\alpha\alpha^r \to 1$, and it can now be proved that also $1 \to \alpha\alpha^l$ and $1 \to \alpha^r\alpha$. Indeed, from $\alpha^l\alpha \to 1$ we infer that

$$1 = 1^r \to (\alpha^l\alpha)^r \to \alpha^r\alpha^{lr} \to \alpha^r\alpha$$

and similarly that $a \to \alpha\alpha^l$. [5]

The hypothetical scenario I have described does not produce arbitrary pregroups, but what mathematicians would describe as *pregroups freely generated by partially ordered sets.* I can only suppose that the conjectured historical process arose simultaneously in modern Western European languages, as did the introduction of articles and the erosion of many inflections. What is surprising is that the elegant mathematical concept of pregroups (other than partially ordered groups) seems to have escaped the attention of mathematicians.

On second thought, I have a nagging feeling that I may have overestimated the innate mathematical ability of natural language speakers. In particular, I wonder whether mathematically untutored minds can really grasp the contravariance of adjunction? I had invoked this contravariance to infer from the postulate $\hat{\mathbf{o}} \to \mathbf{o}$ that $\mathbf{o}^l \to \hat{\mathbf{o}}^l$ and hence $\hat{\mathbf{o}}^{ll}\mathbf{o}^l \to 1$. Now Anne Preller [26] made the brilliant observation: an alternative type assignment can avoid the postulate $\hat{\mathbf{o}} \to \mathbf{o}$ and the necessity of invoking the contravariance of adjunction, with the payoff that forbidden contractions will no longer arise and that my reformulated island constraints [17] are no longer needed!

[5] I am indebted to Michael Barr for pointing out this simple derivation in [16].

Appendix II.
Towards Capulet Semantics.

Let me now turn to the question of semantics. One thing people liked about the original syntactic calculus [12] was that it afforded an easy transition from syntax to Curry-Montague semantics. Take the simple sentence

$$\begin{array}{c} \textit{John loves Jane} \\ \mathbf{n}\ (\mathbf{n} \setminus \mathbf{s})/\mathbf{n}\ \mathbf{n} \end{array}$$

with its original type assignment. If $\mathbf{n}$ is interpreted as the set of persons and $\mathbf{s}$ as the set of truth-values, *loves* should denote a function from $\mathbf{n} \times \mathbf{n}$ to $\mathbf{s}$, more precisely an element of

$$\mathbf{s}^{\mathbf{n}\times\mathbf{n}} \simeq (\mathbf{s}^{\mathbf{n}})^{\mathbf{n}}$$

in the category of sets. Not that this observation tells us anything about the nature of love, still it helps us in translating English into the language of set theory. For this purpose, we need to know nothing about the category of sets, except that it is "cartesian closed", a notion introduced by Bill Lawvere.

Now the arrows $I \to A$ from the terminal object I into and object A (also called "global elements" of A) in a cartesian closed category can be described as *combinators* (Curry) or as *lambda terms* (Church). According to the so-called "Curry-Howard isomorphism", they may also be viewed as proofs in the positive intuitionistic propositional calculus, which may be obtained from the syntactic calculus by introducing Gentzen's three structural rules of *contraction, weakening* and *interchange.* There is thus a possible transition from derivations in the syntactic calculus to deductions in the positive intuitionistic propositional calculus. This transition may be viewed as a functor from a residuated category to a cartesian closed one. It has been exploited most prominently by Montague, though without the language of category theory, and is seen by many people as a valid approach to semantics.

One reason for criticizing the pregroup approach is that the corresponding logical system, known as "compact bilinear logic", is not a conservative extension of the syntactic calculus. For example, $(ab)/c$ and $a(b/c)$ both become abc^l in a pregroup. So how should we interpret abc^l in a cartesian closed category? Well, it has two possible interpretations, namely $c^{a\times b} \simeq (c^b)^a$ and $a \times c^b$, and people dedicated to Montague semantics will have to allow both. I presume that people objecting to this proposal would already have made the choice between $(ab)/c$ and $a(b/c)$ at the syntactic level. I suggest that they can make the same choice at the syntactic level of a pregroup grammar by introducing brackets to distinguish between $[ab]c^l$ and $a[bc^l]$. Such brackets act like punctuation marks in written English; they are not to be taken as elements of the pregroup, as is Mario Fadda's symbol β in [8].

Let me take this opportunity to point out again that I am not persuaded by any semantics that ignores the meanings offered by lexicographers, for example, that the primary meaning of *"uncle"* is "male (spouse of) sibling of parent".[6]

Appendix III.
Another look at crossed dependencies in Dutch.

Dutch, like Swiss German, exhibits the phenomenon of *crossed dependencies.* While, technically speaking, the standard Dutch examples can be handled with context-free grammars (see the lively debate in Savitch et al. [27][7]), such grammars, including pregroup ones, will ignore the connection between a verb infinitive and its (accusative) subject. The following comparison between subordinate clauses in English, German and Dutch will illustrate the problem. I have taken the liberty to replace German and Dutch words by English ones.

English: *that he will let her see Peter kiss Mary*
$$(\bar{\mathbf{s}}\mathbf{s}^l)\pi_3(\pi_3^r\mathbf{s}_1\mathbf{i}^l)(\mathbf{i}\mathbf{i}^l\mathbf{o}_1^l)\mathbf{o}_1(\mathbf{i}\mathbf{i}^l\mathbf{o}_2^l)\mathbf{o}_2(\mathbf{i}\mathbf{o}_3^l)\mathbf{o}_3 \quad \to \bar{\mathbf{s}}$$

Here subscripts have been placed on the object types to indicate their connection to the verbs whose objects they represent. The complete subordinate clause with complementizer has been given type $\bar{\mathbf{s}}$.

German: *that he her Peter Mary kiss see let will*
$$(\bar{\mathbf{s}}\mathbf{t}^l)\pi_3\mathbf{o}_1\mathbf{o}_2\mathbf{o}_3(\mathbf{o}_3^r\mathbf{i})(\mathbf{i}^r\mathbf{o}_2^r\mathbf{i})(\mathbf{i}^r\mathbf{o}_1^r\mathbf{i})(\mathbf{i}^r\pi_3^r\mathbf{t}_1) \quad \to \bar{\mathbf{s}}$$

Here the type $\mathbf{t}_1 \to \mathbf{t}$ distinguishes a subordinate clause (prior to adding a complementizer) from a declarative sentence.

Dutch: *that he her Peter Mary will let see kiss*
$$(\bar{\mathbf{s}}\mathbf{t}^l\langle)\pi_3\mathbf{o}_1\mathbf{o}_2\mathbf{o}_3(\mathbf{t}_1\rangle\pi_3^r\mathbf{i}^l)(\mathbf{i}\mathbf{o}_1^r\mathbf{i}^l)(\mathbf{i}\mathbf{o}_2^r\mathbf{i}^l)(\mathbf{i}\mathbf{o}_3^r) \quad \to \bar{\mathbf{s}}$$

Here I have inserted the braces $\langle$ and $\rangle$ into the types of the complementizer *that* and the finite modal verb-form *will*, respectively. To calculate the type of the complete subordinate clause, we require only the following *inversion rule*:

[6]Etymologically, the word "uncle" is derived from Latin "avunculus", meaning "little maternal grandfather". It replaces now obsolete terms that distinguished between the mother's and the father's brother.

[7]I wish to thank Anne Preller for bringing this useful anthology to my attention.

$$\langle x_1...x_n \rangle \to x_n...x_1.$$

Based on the limited sample at hand, I am not sure whether the x_i here should be restricted to be simple types, basic types or types of nominal expressions. At one stage of the above calculation, we arrive at the type

$$\bar{\mathbf{s}}\mathbf{t}^l \langle \pi_3 \mathbf{o}_1 \mathbf{o}_2 \mathbf{o}_3 \mathbf{t}_1 \rangle \pi_3^r \mathbf{i}^l \to \bar{\mathbf{s}}\mathbf{t}^l \mathbf{t}_1 \mathbf{o}_3 \mathbf{o}_2 \mathbf{o}_1 \pi_3 \pi_3^r \mathbf{i}^l \to \bar{\mathbf{s}} \mathbf{o}_3 \mathbf{o}_2 \mathbf{o}_1 \mathbf{i}^l$$

Before contraction this involves nine simple types (not counting the braces), which I would like to identify with Miller's [21] *chunks of information*, thus reaching the upper limit of people's capacity to process information.

Unfortunately, the free pregroup, adorned with braces and subject to the inversion rule, is no longer a free pregroup. Alternatively, one can incorporate diamonds and boxes into the algebra, as in Moortgat [23], who requires four rewrite rules however.

Note added in proof

Since writing the above in 2004, I have come across some use for triple adjoints in English. For example, in the question

With whom did she go?

it is convenient to assign to the preposition *with* the type $\mathbf{q}\mathbf{o}^{lll}\mathbf{q}^l$.

Acknowledgements. I wish to thank Claudia Casadio and Anne Preller for stimulating discussions, and Michael Moortgat and Dick Oehrle for keeping me on my toes by bringing up challenging examples and suggesting an interesting alternative approach.

The author acknowledges support from the Social Sciences and Humanities Research Council of Canada.

References

[1] Bargelli, D., 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, Berlin 2001, pp. 62–78.

[2] Bargelli, D., and J. Lambek, 'An algebraic approach to Arabic sentence structure', *Linguistic Analysis* 31 (2001), 301–315.

[3] Bargelli, D., and J. Lambek, 'An algebraic approach to Turkish syntax and morphology', to appear.

[4] Buszkowski, W., 'Sequent systems for compact bilinear logic', *Mathematical Logic Quarterly* 49 (2003), 467–474.

[5] Casadio, C., and J. Lambek, 'An algebraic analysis of clitic pronouns in Italian', in P. de Groote et al. (eds.), *Logical aspects of computational linguistics*, Springer LNAI 2099, Berlin 2001, pp. 110–124.

[6] Casadio, C., and J. Lambek, 'A computational algebraic approach to Latin grammar', *Research on Language and Computation* 3 (2005), 45–60.

[7] Chomsky, N., *Syntactic structures*, Mouton, The Hague, 1957.

[8] Fadda, M., 'Toward flexible pregroup grammars', in V.M. Abrusci et al. (eds.), *New prespectives in logic and formal linguistics*, Bulzoni Editore, Rome 2002, pp. 95–112.

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

[10] Kanazawa, M., 'The Lambek calculus enriched with additional connectives', *Journal of Logic, Language and Information* 1 (1992), 141–171.

[11] Kiślak-Malinowska, A., 'Pregroups as a tool for typing relative pronouns in Polish', *Proceedings of Categorial Grammars, An efficient tool for Natural Language Processing*, Montpellier, 2004, pp. 114–128.

[12] Lambek, J., 'The mathematics of sentence structure', *The American Mathematical Monthly* 65 (1958), 154–169.

[13] Lambek, J., 'On the calculus of syntactic types', in R.V. Jacobson (ed.), *Structure of language and its mathematical aspects*, AMS Symposia on applied mathematics 12, Providence 1961, pp. 166–178.

[14] Lambek, J., 'Some Galois connections in elementary number theory', *J.Number Theory* 47 (1994), 371–177.

[15] Lambek, J., 'Some lattice models of bilinear logic', *Algebra Universalis* 34 (1995), 541–550.

[16] Lambek J., 'Type grammar revisited', in A. Lecomte et al. (eds.) *Logical Aspects of Computational Linguistics*, Springer LNAI 1582, 1999, pp. 1–27.

[17] Lambek, J., 'A computational algebraic approach to English grammar', *Syntax* 7 (2004), 128–147.

[18] Lambek, J., 'Invisible endings of English adjectives and nouns', to appear.

[19] Lambek, J., and A. Preller, 'An algebraic approach to the German noun phrase', *Linguistic Analysis* 31 (2001), 291–300.

[20] Lambek, J. and A. Preller, 'An algebraic approach to the German sentence', *Linguistic Analysis* 31 (2001), 270–290.

[21] Miller, G.A., 'The magical number seven plus or minus two: some limits on our capacity to process information', *Psychological Review* 63, (1956), 81–97.

[22] Moortgat, M., 'Categorial type logics', in J. van Benthem et al. (eds.), *Handbook of logic and language*, Elsevier, Amsterdam 1997, pp. 93–177.

[23] Moortgat, M., 'Labelled deduction in the composition of form and meaning ', in U. Keyle and H.J. Ohlbach (eds.), *Logic, Language and reasoning, Essays in honour of Dov Gabbay*, Kluwer Academic Publishers, Dordrecht 1999, pp. 377–399.

[24] Moortgat, M., and D. Oehrle, 'Pregroups and type-logical grammar: searching for convergence', in P. Scott et al. (eds.) *Language and grammar: Studies in mathematical linguistics and natural language*, Center for the study of language and information, Stanford 2004, pp. 141–160.

[25] Peirce, C.S., 'The logic of relatives', *The Monist* 7 (1897), 161–217.

[26] Preller, A., 'Pregroups meet constraints on transformations', to appear.
[27] Savitch, W.J., et al. (eds.), *The formal complexity of natural language*, Reidel Publishing Co., Dordrecht, 1987.
[28] Sheiber, S., 'Evidence against the context-freeness of natural languages', *Linguistics and Philosophy* 8 (1985), 333–343.
[29] Twain, M., 'The horrors of the German Language', in *Mark Twain's speeches*, Harper and Row, New York 1910, pp. 45–51.

Joachim Lambek
Department of Mathematics and Statistics
Mc Gill University
Montreal,Quebec
H3A 2K6, Canada