

Lambek grammars based on pregroups

Wojciech Buszkowski
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Poznań Poland
e-mail: buszko@amu.edu.pl

Logical Aspects of Computational Linguistics, LNAI 2099: 95–109,
2001

Abstract

Lambek [13] introduces pregroups as a new framework for syntactic structure. In this paper we prove some new theorems on pregroups and study grammars based on the calculus of free pregroups. We prove that these grammars are equivalent to context-free grammars. We also discuss the relation of pregroups to the Lambek calculus.

1 Introduction and preliminaries

A *pregroup* is a structure $\mathcal{G} = (G, \leq, \cdot, l, r, 1)$ such that $(G, \leq, \cdot, 1)$ is a partially ordered monoid, and l, r are unary operations on G , satisfying the inequalities:

$$\text{(PRE)} \quad a^l a \leq 1 \leq a a^l \text{ and } a a^r \leq a^r a,$$

for all $a \in G$. The elements a^l and a^r are called *the left adjoint* and *the right adjoint*, respectively, of a . Recall that a partially ordered monoid is a monoid (i.e. a semigroup with unit), satisfying the monotony conditions with respect to the partial ordering \leq : if $a \leq b$, then $ca \leq cb$ and $ac \leq bc$, for all elements a, b, c .

The notion of a pregroup, introduced in Lambek [13], is related to the notion of a residuated monoid, known from the theory of partially ordered algebraic systems [8]. A *residuated monoid* is a structure $\mathcal{G} = (G, \leq, \cdot, \backslash, /, 1)$ such that $(G, \leq, \cdot, 1)$ is a partially ordered monoid, and $\backslash, /$ are binary operations on G , fulfilling the equivalences:

$$\text{(RES)} \quad ab \leq c \text{ iff } b \leq a \backslash c \text{ iff } a \leq c / b,$$

for all $a, b, c \in G$. It suffices to assume that (G, \leq) is a poset, and $(G, \cdot, 1)$ is a monoid, since the monotony conditions can be derived, using (RES). Residuated monoids are most general algebraic frames for the Lambek calculus with

empty antecedents: the sequents provable in this calculus express precisely the inequalities valid in all frames of that kind [4].

Now, in any pregroup one can define $a \setminus b = a^r b$ and $a / b = a b^l$, and the defined operations satisfy (RES). Accordingly, every pregroup can be expanded (by definitions) to a residuated monoid. Then, all sequents provable in the Lambek calculus with empty antecedents are valid in all pregroups. As observed by Lambek, the converse does not hold: $(a \cdot b) / c = a \cdot (b / c)$ is true in all pregroups but not all residuated monoids.

First examples of pregroups are partially ordered groups, i.e. structures of the form $(G, \leq, \cdot, ()^{-1}, 1)$ such that $(G, \leq, \cdot, 1)$ is a partially ordered monoid, and $()^{-1}$ is a unary operation on G , satisfying $a^{-1} a = 1 = a a^{-1}$, for all $a \in G$. In a partially ordered group one defines $a^l = a^r = a^{-1}$, and inequalities (PRE) hold true. Conversely, if $a^l = a^r$, for all $a \in G$, then one can define $a^{-1} = a^l$, and the resulting structure is a partially ordered group. We say that a pregroup \mathcal{G} is *proper*, if it is not a group, that means, $a^l \neq a^r$, for some $a \in G$. If \cdot is commutative ($ab = ba$, for all a, b), then the pregroup is not proper (in a group, the converse element a^{-1} is uniquely determined by a). Consequently, proper pregroups must be noncommutative.

Lambek [13] provides only one natural example of a proper pregroup. It consists of all unbounded, monotone functions from the set of integers into itself. For $f(n) = 2n$, one obtains $f^l(n) = [(n+1)/2]$ and $f^r = [n/2]$, hence $f^l \neq f^r$; here $[x]$ denotes the greatest integer $m \leq x$. This example will be further discussed in section 2. This pregroup will be referred to as *the Lambek pregroup*.

Lambek's approach to syntactic structure is based on the notion of a free pregroup, generated by a poset (P, \leq) . A precise definition of this notion will be given in section 3. Roughly speaking, elements of P are treated as atomic types in categorial grammars. For $p \in P$, one forms iterated adjoints $p^{l \dots l}$, $p^{r \dots r}$. Types assigned to expressions are strings $a_1 \dots a_n$ such that a_1, \dots, a_n are atomic types or iterated adjoints. If A, B are types, then $A \leq B$ holds, if this inequality can be derived by (PRE) and the presupposed ordering \leq on P together with antimotony conditions: if $A \leq B$, then $B^l \leq A^l$ and $B^r \leq A^r$, which are true in all pregroups.

Let us illustrate the matter by simple examples, after Lambek [13]. Atomic types π_1, π_2, π_3 are assigned to personal pronouns: π_1 to *I*, π_2 to *you, we, they*, π_3 to *he, she, it*. Types s_1 and s_2 correspond to declarative sentences in the present tense and in the past tense, respectively.

Types of conjugated verb forms are as follows: *go* - types $\pi_1^r s_1$ and $\pi_2^r s_1$, *goes* - type $\pi_3^r s_1$, *went* - types $\pi_k^r s_2$, for $k = 1, 2, 3$. We obtain the following parsing forms:

$$(1) \text{ he goes} - \pi_3(\pi_3^r s_1) \leq s_1,$$

$$(2) \text{ I slept} - \pi_1(\pi_1^r s_2) \leq s_2.$$

Type i is assigned to infinitives. The auxiliary verb *do* in conjugated forms is given types $\pi_k^r s_l i^l$, for $k = 1, 2, 3$ and $l = 1, 2$, which yields:

(3) *he does go* - $\pi_3(\pi_3^r s_1 i^l) i \leq s_1$.

Adverbs (*quietly*) are of type $i^l i$. Lambek proposes to analyse conjugated forms of verbs as modifications of infinitives: *goes* as $C_{31} go$, *going* as $Part go$, *gone* as $Perf go$, where: $C_{kl} = \pi_k^r s_l i^l$, $Part = p_1 i^l$, $Perf = p_2 i^l$. Then, p_1 and p_2 are types of present participle and past participle, respectively, since $p_l i^l i \leq i$, for $l = 1, 2$. Conjugated forms of *be* can be given types $\pi_k^r s_l p_1^l$, for $k = 1, 2, 3$, $l = 1, 2$; so, *am* is of type $\pi_1^r s_1 p_1^l$. We obtain:

(4) *she goes quietly* - $\pi_3(\pi_3^r s_1 i^l) i(i^l i) \leq s_1$,

(5) *I was going quietly* - $\pi_1(\pi_1^r s_2 p_1^l)(p_1 i^l) i(i^l i) \leq s_2$.

To account for transitive verbs Lambek introduces type o of objects (also accusative forms of pronouns) and assigns types $\pi_k^r s_l o^l$, for $k = 1, 2, 3$ and $l = 1, 2$ to conjugated forms of transitive verbs.

(6) *I saw you* - $\pi_1(\pi_1^r s_2 o^l) o \leq s_2$.

Since plural forms of verbs are independent of Person, then one can introduce a common type π of personal pronouns together with inequalities $\pi_k \leq \pi$, for $k = 1, 2, 3$. Then, transitive verbs in past tense can be given a common type $\pi^r s_2 o^l$, and the latter example admits another parsing:

(7) *I saw you* - $\pi_1(\pi^r s_2 o^l) o \leq \pi(\pi^r s_2 o^l) o \leq s_2$.

Modal verbs are given types: $\pi^r s_1 j^l$ for *may, can, will, shall* and $\pi^r s_2 j^l$ for *might, could, would, should*. Here, j is the type of infinitival intransitive verb phrases, and one supposes $i \leq j$.

(8) *I may go* - $\pi_1(\pi^r s_1 j^l) i \leq s_1$,

(9) *I should have loved him* - $\pi_1(\pi^r s_2 j^l)(i p_2^l)(p_2 o^l) o \leq s_2$,

Lambek uses types q_1, q_2 for interrogatives in present tense and past tense, respectively, and a common type q with postulates $q_l \leq q$.

(10) *does he go?* - $(q_1 i^l \pi_3^l) \pi_3 i \leq q_1$.

Type q' is a common type of interrogatives and wh-questions, and one stipulates $q \leq q'$.

(11) *whom does he see?* - $(q' o^l q_1^l)(q_1 i^l \pi_3^l) \pi_3(i o^l) \leq q'$.

The above examples may be enough to give the reader an idea of Lambek's approach. It is interesting to compare this method with the earlier one, based on the Lambek calculus, introduced in Lambek [12]. First, we recall basic notions.

Syntactic types (shortly: types) are formed out of atomic types by means of operation symbols \cdot (product), \backslash (left residuation) and $/$ (right residuation). A, B, C will denote types and Γ, Δ finite (possibly empty) strings of types.

Sequents are of the form $\Gamma \vdash A$. The calculus **L1** (the Lambek calculus with empty antecedents) admits the following axioms and inference rules:

$$\begin{aligned}
& (\text{Ax}) A \vdash A, \\
& (\text{L}\cdot) \frac{\Gamma, A, B, \Gamma' \vdash C}{\Gamma, A \cdot B, \Gamma' \vdash C}, \quad (\text{R}\cdot) \frac{\Gamma \vdash A; \Delta \vdash B}{\Gamma, \Delta \vdash A \cdot B}, \\
& (\text{L}\setminus) \frac{\Gamma, B, \Gamma' \vdash C; \Delta \vdash A}{\Gamma, \Delta, A \setminus B, \Gamma' \vdash C}, \quad (\text{R}\setminus) \frac{A, \Gamma \vdash B}{\Gamma \vdash A \setminus B}, \\
& (\text{L}/) \frac{\Gamma, A, \Gamma' \vdash C; \Delta \vdash B}{\Gamma, A/B, \Delta, \Gamma' \vdash C}, \quad (\text{R}/) \frac{\Gamma, B \vdash A}{\Gamma \vdash A/B}, \\
& (\text{CUT}) \frac{\Gamma, A, \Gamma' \vdash B; \Delta \vdash A}{\Gamma, \Delta, \Gamma' \vdash B}.
\end{aligned}$$

The original calculus **L** (from [12]) results from restricting (R\setminus) and (R/) to nonempty Γ . **L1** is a natural strengthening of the original calculus. As shown by Abrusci [1], **L1** is a conservative fragment of classical noncommutative linear logics. Both calculi admit cut elimination, but (CUT) is necessary for axiomatic extensions of these systems. In models, \vdash is interpreted by \leq , and $\vdash A$ is true, if $1 \leq \mu(A)$ ($\mu(A)$ is the element of the model which interprets type A).

Let R be a calculus of types. An R -grammar is a quadruple (V, I, s, R) such that V is a nonempty, finite lexicon (alphabet), I is a mapping which assigns a finite set of types to each word from V , and s is a distinguished atomic type. Different calculi R determine different classes of categorial grammars. **L**-grammars and **L1**-grammars are two kinds of Lambek categorial grammars. Classical categorial grammars are based on the calculus **AB**, of Ajdukiewicz [2] and Bar-Hillel [3], which admits product-free types only and can be axiomatized by (Ax), (L\setminus) and (L/). Then, **AB** is a subsystem of **L**. One says that the R -grammar *assigns* type A to string $v_1 \dots v_n$, for $v_i \in V$, if there exist types $A_i \in I(v_i)$, for $i = 1, \dots, n$, such that sequent $A_1 \dots A_n \vdash A$ is provable in R . The *language* of this grammar is defined as the set of all nonempty strings on V which are assigned type s . As shown by Pentus [14], the languages of **L**-grammars are precisely the context-free languages (not containing the empty string), and the same is true for **L1**-grammars [6] and **AB**-grammars [3].

It has been observed by Kiślak [11] that parsing examples provided in Lambek [13] can also be accomplished in **L**-grammars; precisely, in some cases one must add to **L** postulates $A \vdash B$, for some atomic A, B . Actually, the types assigned by Lambek are translations of syntactic types according to the rules: $A \setminus B = A^r B$, $A/B = AB^l$, using equalities $(AB)^l = B^l A^l$ and $(AB)^r = B^r A^r$, which are valid in pregroups (see section 2). For (1)-(11), we obtain the following sequents, derivable in **L** ((1)-(10) are even derivable in **AB!**).

- (1') $\pi_3, \pi_3 \setminus s_1 \vdash s_1$,
- (2') $\pi_1, \pi_1 \setminus s_2 \vdash s_2$,
- (3') $\pi_3, (\pi_3 \setminus s_1)/i, i \vdash s_1$,

- (4') $\pi_3, (\pi_3 \setminus s_1) / i, i, i \setminus i \vdash s_1,$
- (5') $\pi_1, (\pi_1 \setminus s_2) / p_1, p_1 / i, i, i \setminus i \vdash s_2,$
- (6') $\pi_1, (\pi_1 \setminus s_2) / o, o \vdash s_2,$
- (7') $\pi_1, (\pi \setminus s_2) / o, o \vdash s_2$ assuming $\pi_1 \vdash \pi,$
- (8') $\pi_1, (\pi \setminus s_1) / j, i \vdash s_1$ assuming $i \vdash j,$
- (9') $\pi_1, (\pi \setminus s_2) / j, i / p_2, p_2 / o, o \vdash s_2$ under assumptions from (7'), (8'),
- (10') $(q_1 / i) / \pi_3, \pi_3, i \vdash q_1,$
- (11') $q' / (q_1 / o), (q_1 / i) / \pi_3, \pi_3, i / o \vdash q'.$

This observation might suggest that **L1** is equivalent to the calculus of pregroups for sequents $A_1 \dots A_n \vdash A$ such that A_1, \dots, A_n are product-free and A is atomic. This conjecture is false. The sequent $(p / ((p/p)/p)) / p \vdash p$ with p atomic, is not provable in **L1** (use cut-free derivations and the fact that $\vdash p$ is not provable), but its translation is $pppp^l p^l \leq p$, which is valid in pregroups. In section 3, we show that the calculus of pregroups is equivalent to **AB**, hence also to **L** and **L1**, for sequents $A_1, \dots, A_n \vdash A$ such that A_1, \dots, A_n are product-free types of order not greater than 1 and A is atomic. Recall that *the order* of a product-free type A ($o(A)$) is defined as follows: (i) $o(A) = 0$, for atomic A , (ii) $o(A/B) = o(B \setminus A) = \max(o(A), o(B) + 1)$ (see [7]). Sequents (1')-(10') are of that form, whereas in (11') the left-most type is of order 2. Yet, the latter sequent is provable in **L**. This leads us to the open question: are there any ‘linguistic constructions’ which can be parsed by means of pregroups but not the Lambek calculus? In other words: do any syntactic structures in natural language require types of order greater than 1 such that the resulting sequents are valid in pregroups but not provable in **L1**? Evidently, this problem cannot be solved by purely mathematical methods.

We briefly describe further contents. In section 2, we prove some theorems on pregroups. In particular, we prove: (i) no fully ordered pregroup is proper, (ii) no finite pregroup is proper, (iii) every pregroup of all unbounded, monotone functions on a fully ordered set (P, \leq) , P infinite, is isomorphic to the Lambek pregroup. In section 3, we define grammars based on the calculus of free pregroups and prove that the languages of these grammars are precisely the context-free languages. It is worthy of noticing that this proof is easier than the proofs of analogous theorems for **L** and **L1** as well as the nonassociative Lambek calculus [5, 10]. In general, pregroups are definitely easier to handle than residuated semigroups and monoids; on the other hand, the latter are abundant in natural linguistic frames, which is not the case for the former.

2 Pregroups

The aim of this section is to throw more light on the notion of a pregroup by establishing some basic properties of these structures. We are motivated by

Lambek's confession that he does not know any 'natural model' except for the pregroup of unbounded, monotone functions on the set of integers (and p.o. groups, of course). Our results can help to understand the reasons.

The following conditions are valid in pregroups:

$$1^l = 1 = 1^r, a^{lr} = a = a^{rl}, (ab)^l = b^l a^l, (ab)^r = b^r a^r, \\ aa^l a = a, aa^r a = a, \text{ if } a \leq b \text{ then } b^l \leq a^l \text{ and } b^r \leq a^r.$$

Proofs are elementary. We consider one case. By (PRE),

$$a \leq a^{lr} a^l a \leq a^{lr} \text{ and } a^{lr} \leq aa^l a^{lr} \leq a,$$

which yields $a^{lr} = a$. One can also use:

Proposition 1 *In any partially ordered monoid, for every element a , there exists at most one element b such that $ba \leq 1 \leq ab$, and there exists at most one element c such that $ac \leq 1 \leq ca$.*

PROOF. If $ba \leq 1 \leq ab$ and $b'a \leq 1 \leq ab'$, then $b \leq bab' \leq b'$ and $b' \leq b'ab \leq b$, hence $b = b'$. \square

Let (P, \leq) be a poset. A function $f : P \rightarrow P$ is said to be *monotone*, if $x \leq y$ entails $f(x) \leq f(y)$. For functions f, g , one defines $(f \circ g)(x) = f(g(x))$, and $f \leq g$ iff, for all $x \in P$, $f(x) \leq g(x)$. I denotes the identity function: $I(x) = x$. The set $M(P)$, of all monotone functions from P into P with \leq, \circ and I is a p.o. monoid. A pregroup is called a *pregroup of functions*, if it is a substructure (as a p.o. monoid) of the p.o. monoid $M(P)$, for some poset (P, \leq) .

Proposition 2 *Every pregroup is isomorphic to a pregroup of functions.*

PROOF. Let $(G, \leq, \cdot, l, r, 1)$ be a pregroup. We consider the poset (G, \leq) . For $a \in G$, we define $f_a(x) = ax$. Clearly, f_a is a monotone functions from G into G . Further, $f_{ab} = f_a \circ f_b$, $f_1 = I$, and $a \leq b$ iff $f_a \leq f_b$. Then, the mapping $h(a) = f_a$ is an isomorphic embedding of $(G, \leq, \cdot, 1)$ into $M(G)$. We define adjoints: $f_a^l = f_{a^l}$ and $f_a^r = f_{a^r}$. Then, (PRE) hold, and h is the required isomorphism. \square

Proposition 3 *Let F be a pregroup of functions on a poset (P, \leq) . For every $f \in F$, the following equalities are true:*

$$f^l(x) = \min\{y \in P : x \leq f(y)\}, \text{ for all } x \in P, \\ f^r(x) = \max\{y \in P : f(y) \leq x\}, \text{ for all } x \in P.$$

PROOF. Since $f^l(f(x)) \leq x \leq f(f^l(x))$, for all x , then $f^l(x)$ belongs to the set of all y such that $x \leq f(y)$. Let z also belong to this set. Then, $x \leq f(z)$, and consequently, $f^l(x) \leq f^l(f(z)) \leq z$. This proves the first equality. The second one is dual. \square

Conversely, if F is a substructure of $M(P)$, and for every $f \in F$, the functions f^l, f^r , defined above, exist and belong to F , then F is a pregroup of functions. For $f^l(f(x)) = \min\{y : f(x) \leq f(y)\} \leq x$ and $x \leq f(f^l(x))$. The second part of (PRE) can be proved in a similar way.

Corollary 1 *If F is a pregroup of functions on a poset (P, \leq) , then all functions in F are unbounded, that means, $\forall x \exists y (x \leq f(y))$ and $\forall x \exists y (f(y) \leq x)$.*

PROOF. The sets on the right hand side of equalities from proposition 3 must be nonempty. \square

Accordingly, pregroups of functions must consist of unbounded, monotone functions on a poset. The Lambek pregroup F_L consists of all unbounded, monotone functions on (\mathbf{Z}, \leq) , where (\mathbf{Z}, \leq) is the set of integers with the natural ordering. Since the composition of two unbounded functions is unbounded, then the Lambek pregroup is a substructure of $M(\mathbf{Z})$, hence it is a p.o. monoid. For every $f \in F_L$, functions f^l, f^r , defined (!) as in proposition 3, exist and are monotone. We show that f^l is unbounded. Since $f^l(f(x)) \leq x$, then f^l is unbounded in the negative direction. If f^l were bounded in the positive direction, then $f \circ f^l$ would also be bounded, against $x \leq f(f^l(x))$. For f^r , the reasoning is dual. Consequently, $f^l, f^r \in F_L$, hence F_L is a pregroup. For $f(x) = 2x$, one obtains:

$$f^l(x) = \min\{y : x \leq 2y\} = \lceil (x+1)/2 \rceil, \quad f^r(x) = \max\{y : 2y \leq x\} = \lfloor x/2 \rfloor,$$

which yields $f^l \neq f^r$. Thus, the Lambek pregroup is proper.

Lemma 1 *Let F be a pregroup of functions on a poset (P, \leq) , and let a be the first or the last element of (P, \leq) . Then, for all $f \in F$ and $x \in P$, $f(x) = a$ iff $x = a$.*

PROOF. We only consider the case a be the first element. Let $f \in F$. Since f is unbounded, then $f(y) = a$, for some $y \in P$, and consequently, $f(a) = a$, by monotony. Assume $f(x) = a$. Then, $f^r(a) \geq x$, by proposition 3, hence $x = a$ due to the fact $f^r(a) = a$. \square

We prove that the Lambek pregroup is (up to isomorphism) the only pregroup of all unbounded, monotone functions on an infinite, fully ordered set.

Theorem 1 *Let (P, \leq) be a fully ordered set of cardinality greater than 2 such that the family of all unbounded, monotone functions on (P, \leq) is a pregroup. Then, (P, \leq) is isomorphic to (\mathbf{Z}, \leq) .*

PROOF. First, we show that (P, \leq) has no endpoints. Suppose a be the first element of (P, \leq) . By the cardinality assumption, there exist $b, c \in P$ such that $a < b < c$. We define $f(x) = a$, for all $x \leq b$, $f(x) = x$, for all $x > b$. Clearly, f is monotone and unbounded, which contradicts lemma 1, since $f(b) = a$, $b \neq a$. If a is the last element, the reasoning is similar.

Consequently, P is infinite. It suffices to show that, for all $a, b \in P$, if $a < b$, then $\{x \in P : a < x < b\}$ is a finite set. Suppose this set be infinite, for some $a < b$. We consider two cases.

(I) $\{x : a < x < b\}$ is well-ordered by \leq . Then, there exists an infinite chain (x_n) such that $a < x_n < b$ and $x_n < x_{n+1}$, for all $n \geq 0$. We define a function $f : P \mapsto P$ by setting: $f(x) = x$, if $x < a$ or $x > x_n$, for all $n \geq 0$,

and $f(x) = a$, otherwise. Clearly, f is unbounded and monotone. On the other hand, $\{y : f(y) \leq a\}$ is cofinal with (x_n) , hence it contains no maximal element. Then, $f^r(a)$ does not exist, which contradicts the assumptions of the theorem.

(II) $\{x : a < x < b\}$ is not well-ordered by \leq . Then, there exists an infinite chain (x_n) such that $a < x_n < b$ and $x_{n+1} < x_n$, for all $n \geq 0$. We define a function $f : P \mapsto P$, by setting: $f(x) = x$, if $x > b$ or $x < x_n$, for all $n \geq 0$, and $f(x) = b$, otherwise. Again, f is unbounded and monotone. The set $\{y : b \leq f(y)\}$ is cofinal with (x_n) , hence it contains no minimal element. Then, $f^l(b)$ does not exist. \square

Pregroups are closed under direct products. Thus, from the Lambek pregroup one can construct other proper pregroups. The simplest one is $F_L \times F_L$. It is easy to see that $F_L \times F_L$ is isomorphic to the following pregroup of functions on a fully ordered set. This set is the union $Z_1 \cup Z_2$ of two copies of \mathbf{Z} with the natural ordering on Z_1 and Z_2 and, additionally, $x < y$, for all $x \in Z_1, y \in Z_2$. $F_L \times F_L$ can be represented as the family of all functions on this poset whose restrictions to Z_i are unbounded and monotone mappings on Z_i , for $i = 1, 2$. Clearly, this new pregroup consists of some but not all unbounded, monotone mappings from $Z_1 \cup Z_2$ into itself. Other interesting pregroups are countable substructures of the Lambek pregroup (not considered here).

Lemma 2 *For every element a of a pregroup, the following conditions are equivalent: (i) $a^l \leq a^r$, (ii) $aa^l = 1$, (iii) $aa^r = 1$.*

Lemma 3 *For every element a of a pregroup, the following conditions are equivalent: (i) $a^r \leq a^l$, (ii) $a^l a = 1$, (iii) $a^r a = 1$.*

PROOF. We only prove lemma 2. Assume (i). Then, $1 \leq aa^l \leq aa^r \leq 1$, hence (ii) and (iii) hold. Assume (ii). Then, $a^r = a^r aa^l \geq a^l$, hence (i) holds. (iii) entails (i), by a similar argument. \square

An element a fulfilling conditions of lemma 2 (resp. lemma 3) is called *surjective* (resp. *injective*); it is called *bijective*, if $a^l = a^r$. In pregroups of functions, f is surjective (resp. injective) in the sense of pregroups iff f is surjective (resp. injective) as a mapping. For assume $f \circ f^l = I$. Then, $f \circ f^l$ is a surjective mapping, hence f must be a surjective mapping. Assume f be a surjective mapping. Then, $f \circ f^l \circ f = f$ yields $(f \circ f^l)(f(x)) = f(x)$, for all x , and consequently, $f \circ f^l = I$. The second claim is proved in a similar way. Consequently, a pregroup of functions is proper iff not all functions in it are bijective mappings. We have got another evidence that the Lambek pregroup is proper.

Using the above observations and proposition 2, one obtains the following facts, true for every pregroup: (1) if ab is injective, then b is injective, (2) if ab is surjective, then a is surjective, (3) if a is injective (resp. surjective), then a^l and a^r are surjective (resp. injective).

Lemma 4 *If every element of a pregroup is injective or surjective, then every element of this pregroup is bijective.*

PROOF. Assume the antecedent. Suppose a be not injective. Then, $a^l a$ is not injective, by (1), hence $a^l a$ is surjective, and consequently, a^l is surjective, by (2). Then, a is injective, by (3). Accordingly, all elements are injective. By (3), all elements are surjective. \square

Theorem 2 *No fully ordered pregroup is proper.*

PROOF. In a fully ordered pregroup, $a^l \leq a^r$ or $a^r \leq a^l$, for all elements a . Use lemma 4. \square

Theorem 3 *No finite pregroup is proper.*

PROOF. We prove more: if $(G, \leq, \cdot, l, r, 1)$ is a finite pregroup, then \leq is the identity relation, and consequently, this pregroup must be a group. It is known that there are no finite p.o. groups with a nontrivial ordering [8], but the proof for pregroups is different.

In pregroups, $aa^r b \leq b$ and $ba^r a \geq b$. Consequently, for all a, b , there exist x, y such that $ax \leq b$ and $ya \geq b$. Now, let a be a minimal and b be a maximal element of (G, \leq) . There exists x such that $bx \leq a$, hence there exists a minimal element x such that $bx = a$. Let x_1, \dots, x_n be all minimal elements. Then, the set $\{bx_1, \dots, bx_n\}$ must contain all elements x_1, \dots, x_n , hence this set equals $\{x_1, \dots, x_n\}$. As a consequence, bx is minimal, for every minimal element x . A dual argument yields: if x is minimal, and y is maximal, then yx is maximal. Since b is maximal, then all elements bx_i are maximal. Accordingly, all minimal elements are maximal, which proves the theorem. \square

At the end, we show that there exist no proper pregroups of functions on the set of real numbers or the set of rational numbers. First, we observe that every pregroup of functions on a fully ordered set must consist of continuous functions only. The function $f : P \mapsto P$ is *continuous*, if it satisfies the conditions: (i) if $x = \sup\{y : y < x\}$, then $f(x) = \sup\{f(y) : y < x\}$, (ii) if $x = \inf\{y : y > x\}$, then $f(x) = \inf\{f(y) : y > x\}$. Assume f be not continuous. Let $x \in P$ does not fulfil (i). Since f is monotone, then $f(x)$ is an upper bound of $\{f(y) : y < x\}$. Since $f(x)$ is not the least upper bound, then there exists another upper bound $z < f(x)$. Yet, $f^r(z) = \max\{y : f(y) \leq z\} = \max\{y : y < x\}$ does not exist. If x does not fulfil (ii), the reasoning is similar.

Let (P, \leq) be a dense, fully ordered set. Let F be a pregroup of functions on (P, \leq) . We show that all functions in F are injective; so, they are also bijective, by lemma 4. Suppose $f \in F$ be not injective. Then, there are $x, y \in P$ such that $x < y$ and $f(x) = f(y)$. Set $a = f(x)$. By lemma 1, a is not an endpoint of P . Since P is dense, $a = \sup\{z : z < a\}$. We have $f^r(a) \geq y$, but $f^r(z) < x$, for all $z < a$. Consequently, f^r is not continuous, which contradicts the above paragraph.

3 Free pregroups and grammars

For any element a of a pregroup, one defines an element $a^{(n)}$, for $n \in \mathbf{Z}$: $a^{(0)} = a$, $a^{(n+1)} = (a^n)^r$, for $n \geq 0$, $a^{(n-1)} = (a^{(n)})^l$, for $n \leq 0$. Then, the equalities

$a^{(n+1)} = (a^{(n)})^r$ and $a^{(n-1)} = (a^n)^l$ hold true, for all $n \in \mathbf{Z}$. As a consequence of (PRE) and antimonotony conditions, we obtain:

$$a^{(n)}a^{(n+1)} \leq 1 \leq a^{(n+1)}a^{(n)}, \text{ for all } n \in \mathbf{Z}.$$

$$\text{if } a \leq b \text{ then } a^{(2n)} \leq b^{(2n)} \text{ and } b^{(2n+1)} \leq a^{(2n+1)}, \text{ for all } n \in \mathbf{Z}.$$

This motivates the construction of a free pregroup, due to Lambek [13]. Let (P, \leq) be a poset. Elements of P are treated as constant symbols. *Terms* are expressions $a^{(n)}$, for $a \in P$, $n \in \mathbf{Z}$; $a^{(0)}$ is identified with a . *Types* are finite strings of terms.

Quasi-pregroups are defined as pregroups except that \leq need not be anti-symmetrical, that means, \leq is a quasi-ordering. If \leq is a quasi-ordering on X , then one defines $x \sim y$ iff $x \leq y$ and $y \leq x$, for $x, y \in X$. Then, \sim is an equivalence relation, and the quotient relation $[x] \leq [y]$ iff $x \leq y$ is a partial ordering on the quotient set X/\sim . If $(G, \leq, \cdot, l, r, 1)$ is a quasi-pregroup, then \sim is a congruence on this structure (use monotony conditions for \cdot and antimonotony conditions for l, r). We can construct the quotient-structure on G/\sim , by setting: $[x] \cdot [y] = [xy]$, $[x]^l = [x^l]$, $[x]^r = [x^r]$, for $x \in G$. The quotient-structure is a pregroup whose unit element equals $[1]$.

First, we define a quasi-pregroup whose elements are types. For types x, y , $x \cdot y$ is the concatenation of strings x, y , and 1 is the empty string. The adjoints are defined as follows:

$$(a_1^{(n_1)} \dots a_k^{(n_k)})^l = a_k^{(n_k-1)} \dots a_1^{(n_1-1)}, (a_1^{(n_1)} \dots a_k^{(n_k)})^r = a_k^{(n_k+1)} \dots a_1^{(n_1+1)},$$

The quasi-ordering \leq is the reflexive and transitive closure of the relation defined by the following clauses:

$$\text{(CON)} \quad xa^{(n)}a^{(n+1)}y \leq xy \text{ (contraction),}$$

$$\text{(EXP)} \quad xy \leq xa^{(n+1)}a^{(n)}y \text{ (expansion),}$$

$$\text{(IND)} \quad xa^{(2n)}y \leq xb^{(2n)}y \text{ and } xb^{(2n+1)}y \leq xa^{(2n+1)}y \text{ if } a \leq b \text{ in } P \text{ (induced steps),}$$

for all types x, y , $n \in \mathbf{Z}$, and $a, b \in P$,

It is easy to check that this structure is a quasi-pregroup. It is not a pregroup, since $aa^{(1)}a \sim a$ but $aa^{(1)}a \neq a$. The quotient-structure, defined as above, is a pregroup. It is called *the free pregroup* generated by (P, \leq) and denoted $F((P, \leq))$ or $F(P)$, for short.

In what follows, we are mainly concerned with inequalities $x \leq y$ which hold in the quasi-pregroup underlying $F(P)$. Following Lambek, we distinguish two special cases:

$$\text{(GCON)} \quad xa^{(n)}b^{(n+1)}y \leq xy \text{ if either } n \text{ is even and } a \leq b \text{ in } P, \text{ or } n \text{ is odd and } b \leq a \text{ in } P \text{ (generalized contraction),}$$

$$\text{(GEXP)} \quad xy \leq xa^{(n+1)}b^{(n)}y \text{ if either } n \text{ is even and } a \leq b \text{ in } P, \text{ or } n \text{ is odd and } b \leq a \text{ in } P \text{ (generalized expansion).}$$

Clearly, (GCON) can be obtained as (IND) followed by (CON), and (GEXP) as (EXP) followed by (IND). Clearly, (CON) and (EXP) are special cases of (GCON) and (GEXP), respectively. Thus, if $x \leq y$ in the quasi-pregroup, then there exists a derivation $x = x_0 \leq x_1 \leq \dots \leq x_m = y$, $m \geq 0$ such that $x_k \leq x_{k+1}$ is (GEXP), (GCON) or (IND), for all $0 \leq k < m$. The number m is called *the length* of this derivation.

Lemma 5 (*Lambek switching lemma*) *If $x \leq y$ has a derivation of length m , then there exist types z, z' such that $x \leq z$ by generalized contractions or $x = z$, $z \leq z'$ by induced steps or $z = z'$, and $z' \leq y$ by generalized expansions or $z' = y$, and the sum of lengths of these three derivations is not greater than m .*

The key observation is that two adjacent steps of the form (IND)-(GCON) or (GEXP)-(IND) can be interchanged, if the relevant terms do not overlap, and reduced to a single (GCON) or (GEXP), otherwise; also, two adjacent steps of the form (GEXP)-(GCON) can be interchanged, if the relevant terms do not overlap, and reduced to a single induced step, otherwise. Let us demonstrate the latter reduction. There are two possibilities.

(I) $va^{(n)}w \leq va^{(n)}b^{(n+1)}c^{(n)}w \leq vc^{(n)}w$, by (GEXP) and (GCON). Assume n be even. Then, $b \leq c$ and $a \leq b$, which yields $a \leq c$. Clearly, $va^{(n)}w \leq vc^{(n)}w$ is an induced step. Assume n be odd. Then, $c \leq b$ and $b \leq a$, which yields $c \leq a$. Again, these two steps can be reduced to a single induced step.

(II) $vc^{(n+1)}w \leq va^{(n+1)}b^{(n)}c^{(n+1)}w \leq va^{(n+1)}w$, by (GEXP) and (GCON). Assume n be even. Then, $a \leq b$ and $b \leq c$, hence $a \leq c$, and we reason as above. Assume n be odd. Then, $b \leq a$ and $c \leq b$, hence $c \leq a$, and we reason as above.

Now, the lemma can easily be proved by induction on m . In the initial derivation of length m one moves all generalized contractions to the left. If the move is blocked, then the derivation can be replaced by a shorter one, and we apply the induction hypothesis. Next, one moves all generalized expansions to the right.

It follows from the lemma that if $x \leq t$ holds in the quasi-pregroup, and t is a term, then there is a derivation of $x \leq t$ which does not apply (GEXP). Consequently, inequalities $x \leq t$ can be derived by (CON) and (IND) only.

Let (P, \leq) be a finite poset. A *pregroup grammar* based on this poset is defined as a triple (V, I, s) such that V is a finite lexicon (alphabet), I is a mapping which assigns a finite set of types (in the sense of this section) to every $v \in V$, and $s \in P$. We say that this grammar assigns a type x to a string $v_1 \dots v_n$, $v_i \in V$, if there exist types $x_i \in I(v_i)$, $i = 1, \dots, n$, such that $x_1 \dots x_n \leq x$ in the quasi-pregroup underlying $F(P)$. *The language* of this grammar consists of all nonempty strings on V which are assigned type s . Evidently, these notions are analogous to basic notions for categorial grammars (see section 1).

As a consequence of lemma 5, we show that the language of each pregroup grammar must be a context-free language. Fix a pregroup grammar (V, I, s) based on a finite poset (P, \leq) . Let X denote the set of all terms appearing in

types assigned by I to words from V . We define:

$$p = \max\{n \geq 0 : (\exists a \in P)(a^{(n)} \in X \vee a^{(-n)} \in X)\}.$$

By N we denote the set of all terms $a^{(n)}$ such that $a \in P$ and $-p \leq n \leq p$. We show that the set L of all types $t_1 \dots t_n$ such that $n > 0$, $t_i \in N$, for $i = 1, \dots, n$, and $t_1 \dots t_n \leq s$ in the quasi-pregroup underlying $F(P)$ is a context-free language. The context-free grammar which generates L is defined as follows. Both terminal and nonterminal symbols are all terms from N . By lemma 5, $t_1 \dots t_n \leq s$ iff it can be derived by (CON) and (IND) only. Clearly, if $t_1, \dots, t_n \in N$, then all terms in such a derivation are elements of N . Accordingly, production rules of the context-free grammar are of the following form: (i) all rules $t \rightarrow tuv$ such that $t, u, v \in N$ and $tuv \leq t$ is a contraction, (ii) all rules $t \rightarrow uvt$ such that $t, u, v \in N$ and $uvt \leq t$ is a contraction, (iii) all rules $t \rightarrow u$ such that $t, u \in N$ and $u \leq t$ is an induced step.

For any $v \in V$ and any $x \in I(v)$, we create a new symbol v_x . Let V' denote the set of all new symbols. We define a homomorphism f from V' into the set of types, by setting: $f(v_x) = x$. The coimage of L under f is the set:

$$f^{-1}[L] = \{v_{x_1}^1 \dots v_{x_n}^n : n > 0, x_1 \dots x_n \in L\}.$$

Since context-free languages are closed under homomorphic coimages [9], then $f^{-1}[L]$ is a context-free language. We define a homomorphism g from V' into the set of strings on V , by setting: $g(v_x) = v$. Clearly, the language of the pregroup grammar (V, I, s) equals $g[f^{-1}[L]]$. Since context-free languages are also closed under homomorphic images, then the latter language is context-free.

We also prove the converse: every context-free language, not containing the empty string, is the language of some pregroup grammar. We use the classical theorem of Gaifman [3]: every context-free language, not containing the empty string, is the language of an **AB**-grammar whose mapping I uses types of the form $p, p/q, (p/q)/r$ only, where p, q, r are atomic types. These types are of order not greater than 1.

Lemma 6 *Let $A_1, \dots, A_n \vdash p$ be a sequent such that p is atomic, and $o(A_i) \leq 1$, for all $1 \leq i \leq n$. Then, this sequent is derivable in **AB** iff it is derivable in **L1**.*

PROOF. One direction is straightforward, since **AB** is a subsystem of **L1**. We prove the converse direction. Assume the sequent is derivable in **L1**. Then, it has a cut-free derivation in **L1**. By induction on the height of this derivation, we show that it is derivable in **AB**. If it is the axiom $p \vdash p$, the thesis is true. Otherwise, it must be a conclusion of $(L \setminus)$ or $(L /)$. In both cases, the premises are sequents $\Gamma \vdash q$ such that q is atomic and all types in Γ are of order not greater than 1; Γ cannot be empty, since $\vdash q$ has no derivation. By the induction hypothesis, both premises are derivable in **AB**. Since **AB** admits $(L \setminus)$ and $(L /)$, then the conclusion is derivable in **AB**. \square

For every product-free type A , we define a pregroup type $T(A)$, by the following recursion: $T(p) = p$, $T(A \setminus B) = T(A) \cdot T(B)$, $T(A / B) = T(A) \cdot T(B)^l$. Here, we identify atomic types with members of P .

Lemma 7 *Under the assumptions of lemma 6, $A_1 \dots A_n \vdash p$ is derivable in **AB** iff $T(A_1) \dots T(A_n) \leq p$ has a derivation in the sense of free pregroups, by contractions only.*

PROOF. It is well know that $\Gamma \vdash p$ is derivable in **AB** iff Γ can be reduced to p by a finite number of Ajdukiewicz contractions $A, A \setminus B \rightarrow B$ and $A/B, B \rightarrow A$. By (CON), we derive $T(A)T(A \setminus B) \leq T(B)$ and $T(A/B)T(B) \leq T(A)$. Consequently, the implication (\Rightarrow) holds true.

The converse implication is proved by induction on n . Let $n = 1$. Assume $T(A_1) \leq p$ be derivable in free pregroups by contractions. It is easy to see that, if $o(A) \leq 1$, then $T(A)$ is of the form $q_k^r \dots q_1^r p r_1^l \dots r_m^l$, for $k, m \geq 0$ and p, q_i, r_j atomic. Clearly, no contraction can be executed within $T(A)$. So, $T(A_1) \leq p$ must be the trivial case $T(A_1) = A_1 = p$, and $p \vdash p$ is derivable in **AB**. Let $n > 0$. Assume $T(A_1) \dots T(A_n) \leq p$ be derivable in free pregroups by contractions. Consider the first contraction. By the above observation, this contraction cannot be executed within some $T(A_i)$. So, for some $0 \leq i < n$, the relevant terms must be the right-most term of $T(A_i)$ and the left-most term of $T(A_{i+1})$. Assume this contraction is $r^l r \leq 1$. Then, $T(A_i) = q_k^r \dots q_1^r p r_1^l \dots r_m^l r^l$ and $T(A_{i+1}) = r s_1^l \dots s_j^l$. We can represent A_i and A_{i+1} as:

$$A_i = q_k \setminus \dots \setminus q_1 \setminus p / r_1 / \dots / r_m / r,$$

$$A_{i+1} = r / s_1 / \dots / s_j.$$

In A_{i+1} parentheses are associated to the left, and in A_i they are associated to the middle p in any way, since the laws:

$$(A \setminus B) / C \vdash A \setminus (B / C), A \setminus (B / C) \vdash (A \setminus B) / C$$

are derivable in **L1**. We construct the type:

$$B = q_k \setminus \dots \setminus q_1 \setminus p / r_1 / \dots / r_m / s_1 / \dots / s_j.$$

Now, $T(B)$ results from $T(A_i)T(A_{i+1})$ by the contraction, and $A_i A_{i+1} \vdash B$ is derivable in **L1**. The inequality:

$$T(A_1) \dots T(A_{i-1})T(B)T(A_{i+2}) \dots T(A_n) \leq p$$

is derivable by contractions in free pregroups, hence the sequent:

$$A_1 \dots A_{i-1}, B, A_{i+2} \dots A_n \vdash p$$

is derivable in **AB**, by the induction hypothesis. Consequently, the sequent $A_1 \dots A_n \vdash p$ (i.e. our starting sequent) is derivable in **L1**, hence also in **AB**, by lemma 6. If the first contraction is $q q^r \leq 1$, then the reasoning is similar. \square

Let L be a context-free language, not containing the empty string. By Gaifman theorem, there is an **AB**-grammar (V, I, s) , using types of order not greater than 1, which generates this language. We consider the poset $(P, =)$ such that P is the set of atomic types used by this grammar. We define a pregroup

grammar (V, I', s) based on this poset, by setting $I'(v) = \{T(A) : A \in I(v)\}$. Now, $v_1 \dots v_n \in L$ iff there exist types $A_i \in I(v_i)$, for $i = 1, \dots, n$, such that $A_1 \dots A_n \vdash s$ is derivable in **AB** iff there exist pregroup types $T(A_i) \in I'(v_i)$ such that $T(A_1) \dots T(A_n) \leq s$ in free pregroups (lemmas 5, 7) iff $v_1 \dots v_n$ belongs to the language of (V, I', s) . Consequently, L equals the language of the pregroup grammar. We have proved the following theorem.

Theorem 4 *The languages of pregroup grammars are precisely the context-free languages, not containing the empty string.*

References

- [1] V.M. Abrusci, Lambek Calculus, Cyclic Multiplicative-Additive Linear Logic, Noncommutative Multiplicative-Additive Linear Logic: language and sequent calculus, in: V.M. Abrusci and C. Casadio (eds.), *Proofs and Linguistic Categories*, Proc. 1996 Roma Workshop, Bologna, 1996, 21-48.
- [2] K. Ajdukiewicz, Die syntaktische Konnexität, *Studia Philosophica* 1 (1935), 1-27.
- [3] Y. Bar-Hillel, C. Gaifman and E. Shamir, On categorial and phrase structure grammars, *Bull. Res. Council Israel* F 9 (1960), 155-166.
- [4] W. Buszkowski, Completeness results for Lambek Syntactic Calculus, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 32 (1986), 13-28.
- [5] W. Buszkowski, Generative capacity of nonassociative Lambek calculus, *Bull. Polish Academy Scie. Math.* 34 (1986), 507-516.
- [6] W. Buszkowski, Extending Lambek grammars to basic categorial grammars, *Journal of Logic, Language and Information* 5 (1996), 279-295.
- [7] W. Buszkowski, Mathematical linguistics and proof theory, in: J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*, Elsevier, Amsterdam, MIT Press, Cambridge Mass., 1997, 683-736.
- [8] L. Fuchs, *Partially Ordered Algebraic Systems*, Pergamon Press, Oxford, 1963.
- [9] S. Ginsburg, *The Mathematical Theory of Context-Free Languages*, McGraw-Hill, New York, 1966.
- [10] M. Kandulski, The equivalence of nonassociative Lambek categorial grammars and context-free grammars, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 34 (1988), 41-52.
- [11] A. Kiślak, Parsing based on pregroups. Comments on the new Lambek theory of syntactic structure, to appear.

- [12] J. Lambek, The mathematics of sentence structure, *American Mathematical Monthly* 65 (1958), 154-170.
- [13] J. Lambek, Type grammars revisited, in: A. Lecomte, F. Lamarche and G. Perrier (eds.), *Logical Aspects of Computational Linguistics*, LNAI 1582, Springer, Berlin, 1999, 1-27.
- [14] M. Pentus, Lambek grammars are context-free. *Prepublication Series: Mathematical Logic and Theoretical Computer Science* 8, Steklov Mathematical Institute, Moscow, 1992.