

An Adventure into Hungarian Word Order with Cyclic Pregroups

Mehrnoosh Sadrzadeh

ABSTRACT. Pregroups are partially ordered monoids where each element has a left and a right adjoint. They have been developed by Lambek and applied to analysis of syntax of natural languages, e.g. from English and French to Japanese, Arabic, and Persian. In recent joint work with Casadio, we introduced the concept of a cyclic Pregroup to formalize clitic movement in French, Italian, and Persian. Our cyclic meta-rules formalize limited kinds of movements of words within a sentence. This paper is an encounter between cyclic Pregroups and the fragment of Hungarian grammar that deals with change of word order. The movement thereof allows for communicating different kinds of emphasis about the constituents of a sentence. We develop a more general version of our previous rules and use them to derive the new type of the verb from its original type in each case of movement.

1. Introduction

Pregroups are mathematical structures developed by Lambek, as a half-centurium reincarnation of his previous Syntactic Calculus. They formalize regularities of numbers and have applications to Linguistics [12]. In their categorized version, they are compact closed categories and have been applied to physics [10, 2]. In the linguistic world, Pregroups have earned rapid success in describing syntax of natural languages, they have been applied to English, French, Italian, but also to Japanese, Arabic, and Persian, and many more, for references see [7]. Hungarian is missing from the list! We do not claim to attempting to add it in this paper, nonetheless offer a touristic encounter with a small fragment of its syntax. This is the section that deals with change of word order, due to change of focus or emphasis on constituents of the sentence. We attempt to explain and formalize this movement using Pregroup grammars that are equipped with special *cyclic rules*.

In previous work [8], we have argued that certain *cyclic rules*, once limited to particular parts of speech such as verbs, formalize movement in the analysis of syntax of natural languages. We have analyzed how, when applied to the categorial type of the verb, these rules formalize the main patterns of clitic movement, in a particularly appealing and uniform way, in languages of Persian, French and Italian. In this paper we develop more general versions of our previous rules and show how they formalize the dynamic word order of Hungarian. We use the rules to derive the type of the verb after each movement from the original type of the verb in a sentence with basic word order and no emphasis. Our approach to explaining and formalizing movement differs from that suggested in [1, 14, 17].

1991 *Mathematics Subject Classification*. Primary 06F99; Secondary 68T99.

Key words and phrases. Mathematical Linguistics, Pregroups, Hungarian, Syntax, Cyclicity.

I would like to thank Jim Lambek for invaluable discussions. This research was supported in part by EPSRC PDRF Grant EP/E035698/1.

We do not work with *focus* phrases. These may be empty even in the presence of a movement, hence cause ambiguity in the meaning of the sentence. Instead, we introduce the new notion of a *mobilizer*, this is a constituent that causes the movement because the speaker wants to convey extra information about it. In our approach, whenever there is movement, there is a mobilizer, hence the ambiguity gets discarded.

In this paper and in pervious work, we have added the cyclic rules as what Lambek calls meta-rules to the Pregroup dictionaries of natural languages, such that they are only applicable to the type of the verb. It is tempting to want to abstract away and instead add them (in their more general form of this paper or the more restricted form of previous work) as axioms to the axiomatics of a Pregroup, thus make them applicable to all the types rather than just that of the verb. This temptation should be resisted, since as we shall show, such a Pregroup will reduce to a partially ordered group.

To analyze movement, e.g. in the wh-questions of English, Lambek uses the *trace* terminology of Chomsky. He argues that whenever there is a movement in formation of a clause, then there should be an iterated adjoint present in the types involved. Our Pregroup treatment of movement, however original, is not in conflict with Lambek's approach. As shown in section 2.3, our rules are derivable from Pregroup versions of the cyclic rules of non-commutative Linear Logic. These have double adjoints in them and as demonstrated at the end of 2.3, one can apply Lambek's idea to the equations of 2.1 to explain the movement in a beautifully comprehensive manner. We have chosen to work with these rules and not the double adjointed versions, since they express the movements that we are interested in, in a more direct and transparent manner.

2. Pregroups Equipped with Clitic Rules

A Pregroup is a partially ordered monoid where each element has a left and a right adjoint. It is denoted by a tuple $(P, \leq, 1, (-)^l, (-)^r)$ satisfying

$$\begin{aligned} p \leq q &\implies pr \leq qr & \text{and} & & rp \leq rq \\ p^l p &\leq 1 \leq pp^l & \text{and} & & pp^r \leq 1 \leq p^r p \\ p1 &= 1 & \text{and} & & 1p = 1 \end{aligned}$$

It follows that adjoints are unique and that the unit is self adjoint, i.e.

$$1^l = 1 = 1^r$$

The adjoint of multiplication is multiplication of adjoints but in the reverse order, i.e.

$$(pq)^l = q^l p^l \quad \text{and} \quad (pq)^r = q^r p^r$$

An example of a Pregroup from arithmetics is the set of all unbounded monotone maps on integers. The free Pregroup generated over a partially ordered set exists, for more details on these see [13, 12, 4]

Since their introduction by Lambek in [13], Pregroups have been applied to the analysis of syntax of a wide range of natural languages. The natural language applications are based on the following two main definitions.

DEFINITION 2.1. A Pregroup dictionary is a relation, given by

$$\mathcal{D} \subseteq \Sigma \times T(\mathcal{B})$$

where Σ is the set of all words of a natural language, \mathcal{B} is a partially ordered set of grammatical types, and $T(\mathcal{B})$ is the free pregroup generated over \mathcal{B} .

A Pregroup dictionary assigns to each word of a language, a pregroup type, based on the grammatical roles it may take in the language. These types are used to decide whether a string of words constitutes a grammatical phrase or not, via the following definition.

DEFINITION 2.2. Given a set α of designated grammatical types $\alpha \subseteq T(\mathcal{B})$, a string of words $w_1, \dots, w_n \in \Sigma$ is grammatical iff $w_1 \cdot \dots \cdot w_n \leq \sigma$ for $\sigma \in \alpha$.

Examples of elements of α are types standing for declarative sentence and questions. There is a polynomial time algorithm to decide the above procedure, but only for context-free languages. An example of a context-sensitive language is Dutch.

As a rough and non-inclusive example of English grammar, consider the set of types $\{\pi, o, s, q, \bar{q}, i\}$, where π stands for subject, o for object, s for declarative sentence, q for yes-no question, \bar{q} for wh-question, and i for the infinitive of a verb. Consider the following dictionary

$$\{(I, \pi), (\text{him}, o), (\text{saw}, \pi^r s o^l), (\text{did}, q i^l \pi^l), (\text{whom}, \bar{q} o^{ll} q^l), (\text{see}, i o^l)\}$$

According to this dictionary and taking α to include $\{s, q, \bar{q}\}$, the following strings of words are grammatical in English

$$\begin{array}{ccccccc} I & \text{saw} & \text{him} & & & & \\ \pi & (\pi^r s o^l) & o & \leq & s & & \\ \\ \text{Did} & I & \text{see} & \text{him?} & & & \\ q i^l \pi^l & \pi & i o^l & o & \leq & q & \\ \\ \text{Whom} & \text{did} & I & \text{see?} & & & \\ \bar{q} o^{ll} q^l & q i^l \pi^l & \pi & i o^l & \leq & \bar{q} & \end{array}$$

Note the double adjoint o^{ll} in the type of whom; according to Lambek [11], this marks the movement of the object from the end of the sentence to its beginning (in the form of the wh-question word).

To each Pregroup inequality, one can assign a diagram. In the case of inequalities corresponding to grammatical strings of words, these diagrams are referred to as *reduction diagrams*. They depict the grammatical structure of a string of words and are one dimensional alternatives to parse-trees. For example the reduction diagrams of the above sentences are as follows:

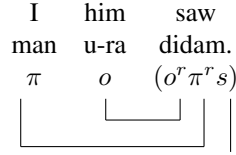
$$\begin{array}{ccccccc} I & \text{saw} & \text{him} & & & & \\ \pi & (\pi^r s o^l) & o & & & & \\ \hline & & & & & & \\ \\ \text{Did} & I & \text{see} & \text{him?} & & & \\ q i^l \pi^l & \pi & i o^l & o & & & \\ \hline & & & & & & \\ \\ \text{Whom} & \text{did} & I & \text{see?} & & & \\ \bar{q} o^{ll} q^l & q i^l \pi^l & \pi & i o^l & & & \\ \hline & & & & & & \end{array}$$

A similar approach is used to formalize grammar of different languages. Reduction diagrams are useful tools in comparing the grammatical structures of different languages.

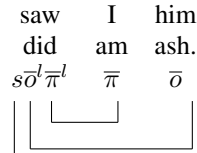
For instance, consider the following two diagrams. They both reduce the sentence "He bought a book from the bookshop", the left hand side is in Persian and Hindi, and the right hand side is in English and French. Comparing these diagrams, without having any knowledge of the corresponding grammars, one can prescribe that English and French are verb-in-the-middle, whereas Persian and Hindi are verb-at-the-end. For more examples, including a poem by Omar Khayam, and details of typing see [15].



2.1. Clitic Movement in Persian. When formalizing the Persian grammar using Pre-groups in [15], I noticed the following phenomenon. The basic word order for a Persian sentence with transitive verb is SOV (Subject-Object-Verb), e.g. in the following sentence.



But there is also another possibility: the subject and object, especially when they are pronouns, may become clitic pronouns and attach themselves to the end of the verb. In so doing, their order of occurrence with regard to each other does not change: the subjective clitic is before the objective one in the clitic sentence, exactly like in the original sentence where the subject is before the object. However, their order does change with regard to the verb. In the clitic sentence the the subject is the closer of the two to the verb, whereas in the normal sentence the object is closer to the verb. For examples, the clitic version of the above sentence is as follows



The over-lined types $\bar{\pi}$, \bar{o} , stand for the clitic versions of the subject and object pronouns. It is rather unpleasant to have to have two different types for the verb, where obviously the clitics have a pattern to their movement. One solution would be to add a rule to the Pregroup dictionary of Persian to be able to derive the clitic type of the verb, i.e. the verb with the clitic pronouns attached to it, from its original type. This rule is as follows

Clitic Rule (1): If $p^r q$ is in the original type of the verb, so is $q\bar{p}^l$.

The over-lined types \bar{p}^l , \bar{p}^r are introduced as a notational convenience to distinguish the clitic pronouns from the non-clitic pronouns or arguments. For any clitic pronoun \bar{p} , we postulate the partial order $\bar{p} \leq p$ to express the fact that a clitic pronoun is also a kind of pronoun. We assume that for all $p, q \in P$, we have $\bar{p}\bar{q} = \bar{p} \bar{q}$. The calculation to derive the clitic type of the verb from its original type in the above example is as follows

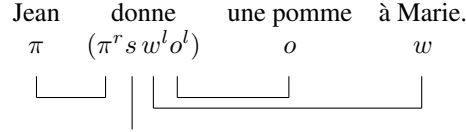
To derive the new type of the verb, we apply rule (1) to its original type $o^r \pi^r s$ by taking q to be s and p to be πo , and noting that $o^r \pi^r = (\pi o)^r$ and $(\pi \bar{o})^l = (\pi \bar{o})^l = \bar{o}^l \pi^l$. Denoting the application of rule (1) by the squiggle arrow, this derivation can be abbreviated to

$$o^r \pi^r s = (\pi o)^r s \rightsquigarrow^1 s(\pi \bar{o})^l = s(\pi \bar{o})^l = s \bar{o}^l \pi^l$$

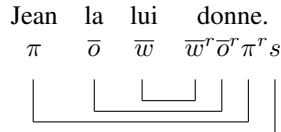
Mention of the word *original* in the rule is to prevent the repetition of its application, so that one cannot apply the rule over and over to the types resulting from the application of the rule itself. If this was allowed, then non-grammatical types would be generated. The bars are to prevent formation of the type (of the verb) that allows the non-clitic pronouns to also appear after the verb, again resulting in ungrammatical sentences. The mention of the word *in* is to formalize partial clitic movements, when only one of the subject or object becomes a clitic. These partial movements are not very common in Persian, so we will not discuss them here.

2.2. Clitic Movement in French. After discussions with Casadio, I came to realize that a similar phenomenon happens in many other languages, including French and Italian. We started to work on a similar formalization for French and Italian, leading to joint work [8]. Interestingly, it turned out that the converse (up to the bars) of the above rule is needed for the clitic movement in French and the movement in Italian needs both rules.

Consider a bi-transitive sentence in French



The clitic movement for this sentence has the following pattern



To be able to derive the clitic type of the verb from its original type, we add the following rule to the pregroup dictionary of French

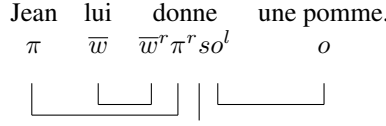
Clitic Rule (2): If qp^l is in the *original* type of the verb, so is $\bar{p}^r q$.

Note that, up to a bar, this is the converse of the clitic rule for Persian. The calculation for the derivation of the above example is as follows

To derive new type of the verb, we apply rule (2) to its original type $(\pi^r s w^l o^l)$ by taking q to be $\pi^r s$ and p to be ow , and noting that $w^l o^l = (ow)^l$ and $(\bar{o} \bar{w})^r = (\bar{o} \bar{w})^r = \bar{w}^r \bar{o}^r$. This derivation can be abbreviated to

$$\pi^r s w^l o^l = \pi^r s (ow)^l \rightsquigarrow^2 (\bar{o} \bar{w})^r \pi^r s = (\bar{o} \bar{w})^r \pi^r s = \bar{w}^r \bar{o}^r \pi^r s$$

Clitics can have a partial movement, i.e. it is possible that only one of the objects moves to before the verb where as the other one stays at its original position after the verb. An example of a sentence with a partial clitic movement is "Jean lui donne une pomme.", which we type as follows



To derive the type of the verb, the same clitic rule is applied, but this time to a subtype of the type of the original verb, via the usage of the word *in*, as follows

Apply rule (2) the subtype $\pi^r s w^l$ of the original type of the verb and turn it into $\bar{w}^r \pi^r s$. Then attach its to the left over o^l and obtain the desired type $\bar{w}^r \pi^r s o^l$.

For more examples in Persian, French, and also Italian, including some exceptions, see [8].

2.3. Cyclic Pregroups and Groups. To formalize clitic movement, we have added rules to the Pregroup dictionaries of languages. It is tempting to want to instead add them as the following axioms to a Pregroup

$$(1) p^r q \leq q p^l \qquad (2) q p^l \leq p^r q$$

In so doing, these become applicable to all the types within the Pregroup, rather than just to the type of the verb as we have been doing so far.

One can then define the notions of a *left cyclic Pregroup*, as a Pregroup with axiom (1), a *right cyclic Pregroup*, as a Pregroup with axiom (2), and a *cyclic Pregroup*, as one with both. The free models in each case would be generated over the already existing partial order \mathcal{B} plus the new type \bar{p} and the new order $p \leq \bar{p}$. However, addition of either of these axioms makes the Pregroup reduce to a partially ordered group, which is an example of a Pregroup where for all $p \in P$ we have $p^l = p^r$. To see this, consider axiom (1), take $q = 1$, we obtain $p^r \leq p^l$ for all $p \in P$, apply left adjoints to both sides and obtain $p^{ll} \leq p$. Here take $p = w^r$ for some $w \in P$ and obtain $w^l \leq w^r$. Now since we have $p^r \leq p^l$ for all p , we obtain $w^l = w^r$. The argument for axiom (2) is similar.

Interestingly, it turns out that the above axioms are derivable from the translation of Abrusci and Yetter's cyclic rules for non-commutative Linear Logic [3, 16]. These are as follows

$$(2.1) \quad (\text{left}) \quad ab \leq ba^{ll} \qquad (\text{right}) \quad ab \leq b^{rr}a$$

To see this, consider Axiom (1), it is derivable from the left axiom by taking a to be p^r and b to be q , since $(p^r)^{ll} = p^l$. Axiom (2) is derivable from the right axiom by taking a to be q , and b to be p^l , since $(p^l)^{rr} = p^r$.

Lambek argues that whenever there is a movement in the formation of the sentence, then iterated adjoints appear in the involved types. This idea exposes itself in a beautiful way in the above equations. These basically say that the juxtaposition of types ab can also be juxtaposed as ba , but in so doing, the moved type is marked with a double adjoint. Hence, the left equation $ab \leq ba^{ll}$ describes the case when the type a moves from the left to the right of the type b , that is why it is annotated with the double adjoint ll . The right equation $ab \leq b^{rr}a$, describes the other possible movement; that it is b that moves from the right side of a to its left, hence it is annotated with the double adjoint rr .

2.4. An Example from Number Theory. Lambek has shown that the set of all monotone maps on integers, i.e. $\mathbb{Z}^{\mathbb{Z}}$ is a Pregroup under function composition and the natural order. This example has been referred to as the Lambek Pregroup. He discusses in detail the example $f(n) = 2n$ by computing the following

$$f^r(n) = \lfloor n/2 \rfloor \quad \text{and} \quad f^l(n) = \lfloor (n+1)/2 \rfloor$$

where

$\lfloor x \rfloor$ is the biggest integer smaller than x .

We use this to build an example for a left cyclic Pregroup. For all $n \in \mathbb{Z}$, we have that

$$\lfloor n/2 \rfloor \leq \lfloor (n+1)/2 \rfloor$$

From this and for $k \in \mathbb{Z}^+$ it follows that

$$k\lfloor n/2 \rfloor \leq k\lfloor (n+1)/2 \rfloor \quad \text{and} \quad \lfloor (kn)/2 \rfloor \leq k\lfloor n/2 \rfloor$$

By transitivity from the above two it follows that

$$\lfloor (kn)/2 \rfloor \leq k\lfloor (n+1)/2 \rfloor$$

So we have shown that

$$f^r \circ g \leq g \circ f^l$$

for

$$g(n) = kn \quad \text{and} \quad f(n) = 2n$$

In other words we have shown that for the above choice of functions from the Lambek Pregroup, the left cyclic axiom holds. Finding a direct example where the right cyclic axiom holds has been more challenging so far.

3. Movement in a Hungarian Sentence

Hungarian belongs to the family of Fino-Ugric languages. This is the largest subgroup of Uralic, spoken by people in northern Eurasia, with Hungary serving as its western limits. Uralic includes more than 30 languages that date as far as the third millennium B.C. It has no relatives in other language families. However (to my delight and surprise), proto-Uralic seems to have been influenced by Indo-Iranian. This similarity is established in [5] and is witnessed by the identity of words, thought to be loan words. The loan dates back to 2-3 millennium B.C. and its assumed place is the banks of river Volga¹. Here are some examples of the loan words

Word	Indo-Iranian	Hungarian
hundred	sata	száz
seven	hapt	hét
horn	sarva	szarv

The words for seven and hundred are still in use in modern Persian.

My references in Hungarian grammar have been [5, 9, 1, 14, 17]. According to the more aged of these references, i.e. [5, 9], word order in a Hungarian sentence is either SOV (Subject, Object, Verb) for predicative sentences or SVO (Subject-Verb-Object) for verbal sentences. They argue that since Hungarian is an agglutinative language, i.e. grammatical relations are expressed by means of affixes, word order does not need much discussion. [9] refers to it as the "analytic" approach to Hungarian grammar and moves to choose a "synthetic" approach which he finds to be more "conspicuous and decisive". [5] discusses it under the section on "Clause Structure", where the closest to the kind of movement that we have in mind is called the "Emphatic Order". The cases discussed therein are rather

¹In second thoughts, this similarity should not be surprising, as it was about the same time that the ancient Aryans (i.e. proto Indo-Iranians) descended from Siberia and on their way south to the plateau of Iran (must have) crossed the Ural mountains and the river Volga.

limited. They only deal with instances where the emphasized constituent moves to the "head of the clause", hence only covering our *multi moves* of subsection 3.2.

My more recent references, i.e. [1, 14, 17], have had more material on word order. They suggest that exactly due to the agglutinative nature of the language, one should not look at word order in Hungarian from the usual SVO or SOV perspective, but rather as a focus-oriented structure. [1] suggests to approach it (Hungarian as well as the whole family of Uralic languages) as a TFV for Topic-Focus-Verb. The other two follow him and work with the "topic + focus + verb + rest" order. The approach of this paper is closer to these more recent ones, but, as we shall see, it is not exactly the same. Our example sentence is from [17], as it had the most comprehensive examples of the three.

The basic word order for a sentence with a transitive verb in Hungarian is as follows

János tegnap elvitt két könyvet Péternek.
 János yesterday took two books to Péternek.

Taking π to stand for the type of subject, w for first object, o for second object, and λ for adverb, we assign the following types² to the constituents of the above sentence

János	tegnap	elvitt	két könyvet	Péternek.
π	λ	$(\lambda^r \pi^r s o^l w^l)$	w	o

The basic order changes according to the information structure of the context in which the sentence is uttered. This is to enable the speaker communicate extra information about a constituent, e.g. emphasize it or make it inclusive. For instance, the speaker might want to emphasize that it was indeed János who took the books, or that it was only two books and nothing more that were taken from one person to another. In any case, we call the constituent about which some extra information is conveyed, a *mobilizer*. This is different from the *focus* phrases of [1] and [14], where the focus can be empty but a movement is still possible. This causes an ambiguity on the existence of a focus and hence on interpreting the meaning of the sentence. From a pure syntactic point of view, if the verb has a prefix, e.g. in *elvitt* and a focus is present, then the prefix becomes a suffix, e.g. *vitt el*. But if there is no prefix then it would be ambiguous as whether there is a focus present or not. Even if not, the word order that corresponds to the existence of the focus is still possible, but will convey a different meaning. Hence, the same sentence may have different possible information structures. Our approach discards this ambiguity: whenever there is a movement, we have a mobilizer. It is based on these mobilizers that we classify and discuss each possible movement.

Our aim is to derive the type of the verb after each possible movement from the type of the verb in the basic sentence, i.e. from $(\lambda^r \pi^r s o^l w^l)$. As the constituents that move are not necessarily clitics and the movement is fairly free, we generalize our previous *clitic rules* in two ways: (1) we remove the bars from the moved types, so that they are not inclusive to clitic pronouns, and (2) we remove the word *original* from the wording, so that the rules may be repetitively applied to intermediate types. This will not produce ungrammatical types anymore, exactly because the movement is not restricted to clitic pronouns. The new rules are as follows

²These types are the most primitive options that serve the purpose of this paper, it is likely that a more sophisticated typing is needed for the Pregroup typing of the full Hungarian grammar.

Movement Rule (1): If $p^r q$ is in the type of the verb, so is qp^l .

Movement Rule (2): If qp^l is in the type of the verb, so is $p^r q$.

Rule (1) describes the movement of a word with type p from before the verb to after it. In the original sentence, this type will cancel out with a type p^r that occurs in the type of the verb. After the movement, it will cancel out with a type p^l in the new type of the verb. Rule (2) describes an opposite movement, i.e. when a word with type p moves from after the verb to before it, hence originally it cancels out with a type p^l within the type of the verb, but after the movement with a p^r .

We distinguish between two types of movement: single and multi. Taking the basic word order as the reference frame, the former is when *only one word* changes its place of occurrence in the sentence, e.g. by moving from after the verb to before it. The latter is when *two or more words* do single moves. The second type of movement and respectively its formalization, is more complex. In fact, it was to accommodate this kind of movement that we had to allow for repetitive applications of the above rules.

3.1. Single Moves. The mobilizer of this type can be the subject or either of the objects. In each case, they will appear right before the verb after the movement. If the mobilizer is the subject, i.e. *János*, then the temporal adverb *yesterday* moves to after the verb, as follows

János	vitt el	tegnap	két könyvet	Péternek.
János	took	yesterday	two books	to Péternek.
π	$(\pi^r so^l w^l \lambda^l)$	λ	w	o

To derive the new type of the verb, we apply rule (1) to the original type $(\lambda^r \pi^r so^l w^l)$, by taking q to be $\pi^r so^l w^l$ and p to be λ .

If the mobilizer is the first object, i.e. *two books*, then it moves to before the verb. The above sentence and its typing change as follows

János	tegnap	két könyvet	vitt el	Péternek.
János	yesterday	two books	took	to Péternek.
π	λ	w	$(w^r \lambda^r \pi^r so^l)$	o

To derive the new type of the verb, we apply rule (2) to the original type $(\lambda^r \pi^r so^l w^l)$, by taking q to be $\lambda^r \pi^r so^l$, and p to be w .

If the mobilizer is the second object *Péternek*, then it moves to before the verb, as follows

János	tegnap	Péternek	vitt el	két könyvet.
János	yesterday	to Péternek	took	two books.
π	λ	o	$(o^r \lambda^r \pi^r sw^l)$	w

To derive the new type of the verb, we apply rule (2) to the subtype $\lambda^r \pi^r so^l$ of the original type, by taking q to be $\lambda^r \pi^r s$ and p to be o .

The verb can also be a mobilizer, e.g. when extra information is conveyed to mark the completion of its corresponding task, as in the task of *taking*. In this case the verb moves to the beginning of the sentence, as follows

Elvitt	János	tegnap	két könyvet	Péternek.
Took	János	yesterday	two books	to Péternek.
$(so^l w^l \lambda^l \pi^l)$	π	λ	w	o

To derive the new type of the verb, we apply rule (1) to the original type $(\lambda^r \pi^r so^l w^l)$, taking p to be $(\pi \lambda)$, and q to be $so^l w^l$. The calculation is abbreviated to

$$\lambda^r \pi^r so^l w^l = (\pi \lambda)^r so^l w^l \rightsquigarrow^1 so^l w^l (\pi \lambda)^l = so^l w^l \lambda^l \pi^l$$

3.2. Multi Moves. Here a different kind of information, rather than just emphasis, is conveyed about some of the constituents. For instance, the information that the object is the sole or not the sole entity on which the verb has taken place. In this case, not only the mobilizer moves to before the verb, but also, together with the verb, they move to the beginning of the sentence.

If the mobilizer is the first object, e.g. we want to say that it was to Péter and no one else that the two books were taken, the sentence becomes as follows

Péternek	vitt el	tegnap	János	két könyvet.
To Péternek	took	yesterday	János	two books.
o	$(o^r sw^l \pi^l \lambda^l)$	λ	π	w

One way to see this as a multi move is: not only Péternek has moved to the beginning of the sentence, but also first tegnap and then János have moved from before the verb to after it, and in so doing have changed their order with regard to each other. The calculation for deriving the new type of the verb reflects the above complications and needs repetitive applications of the rules. It is as follows

Start from the original type of the verb $(\lambda^r \pi^r so^l w^l)$, according to our above description, first Péternek moves to the front, for this we apply rule (2) to the subtype $\lambda^r \pi^r so^l$, take p to be o , and obtain $o^r \lambda^r \pi^r sw^l$. Then tegnap moves to after the verb, for this we apply rule (1) to the subtype $\lambda^r \pi^r sw^l$, take p to be λ , and obtain $o^r \pi^r sw^l \lambda^l$. Finally János moves to after tegnap, for this we apply rule (1) to the subtype $\pi^r sw^l$, take p to be π , and obtain $o^r sw^l \pi^l \lambda^l$.

If the mobilizer is the second object, e.g. if we want to say that two books and nothing else were taken to Péternek, then the sentence becomes as follows

Két könyvet	vitt el	János	tegnap	Péternek.
Two books	took	János	yesterday	to Péternek.
w	$(w^r so^l \lambda^l \pi^l)$	π	λ	o

One way to see this as a multi move is: first két könyvet moves to the beginning of the sentence, then the two words János and tegnap move to after the verb. The corresponding calculation is as follows.

First apply rule (2) to $(\lambda^r \pi^r so^l w^l)$, take p to be w , and obtain $w^r \lambda^r \pi^r so^l$, then apply rule (1) to the subtype $\lambda^r \pi^r so^l$, take p to be $\pi \lambda$, note $\lambda^r \pi^r = (\pi \lambda)^r$, similarly $(\pi \lambda)^l = \lambda^l \pi^l$, and obtain $w^r so^l \lambda^l \pi^l$.

The following is also a possibility, it is uttered when one wants to insist that two books were taken to Péternek but at the same time not exclude the possibility that may be also other things were taken.

Két könyvet	tegnap	elvitt	János	Péternek.
Two books	yesterday	took	János	to Péternek.
w	λ	$(\lambda^r w^r so^l \pi^l)$	π	o

One way to see this as a multi move is: first János moves to after the verb then két könyvet moves to the beginning of the sentence. The calculation is as follows

Starting from the original type of the verb $(\lambda^r \pi^r so^l w^l)$, the calculation to derive these series of moves is as follows: apply rule (1) to the subtype $\pi^r so^l$, take p to be π , and obtain $\lambda^r so^l \pi^l w^l$, then apply rule (2) to the subtype $so^l \pi^l w^l$, take p to be w , and obtain $\lambda^r w^r so^l \pi^l$.

4. Summary

We have added two meta rules to a preliminary Pregroup dictionary of Hungarian. These rules are generalizations of our pervious *cyclic rules* [8]. They are more general since they allow for repetative applications and can also be applied to any word that does not necessarily have a clitic pronoun assigned to it. We have used these rules to formalize and reason about change of word order in Hungarian. We have done so by deriving the new type of the verb after the movement, from its original type before the movement.

Change of word order in Hungarian is due to change of emphasis in a sentence, it occurs when an extra piece of information needs to be communicated about a constituent of the sentence. Some references on Hungarian grammar, e.g. [1, 14, 17] explain this movement by using the notion of *focus*; we introduce the new concept of a *mobilizer*. In some cases there is a movement in the sentence which changes the meaning, but the focus is empty. This causes ambiguity in interpreting the meaning of the sentence. We discard this ambiguity by assigning a mobilizer to each possible movement. This is the constituent that causes the movement and about which extra information is communicated.

It is according to the position of the mobilizer after the movement that a different meaning is assigned to a sentence.

We have shown that the axiomatic version of our rules are derivable from Abrusci and Yetter's cyclic axioms for non-commutative linear logic [3, 16]. We have also reviewed how adding the meta rules as axioms to Pregroups, hence developing the new notion of a *cyclic pregroups*, reduces them to ordered groups. We have found an example of our first rule (left cyclic) within the arithmetics model of a Lambek pregroup.

References

- [1] D. Abondolo, *The Uralic Languages*, Routledge, London, New York, 1998.
- [2] S. Abramsky and B. Coecke, 'A Categorical Semantics of Quantum Protocols', in Proceedings of the 19th Annual IEEE Symposium on *Logic in Computer Science*, 2004.
- [3] M. Abrusci, 'Classical Conservative Extensions of Lambek Calculus', *Studia Logica* **71**, 2002.
- [4] M. Barr, 'On Subgroups of The Lambek Pregroup', *Theory and Application of Categories* **12**, 2004.
- [5] L. Benkő and S. Imre, *The Hungarian Language*, Mouton, The Hague, Paris, 1972.
- [6] W. Buszkowski, 'Lambek grammars based on Pregroups', in *Logical Aspects of Computational Linguistics*, P. de Groote, G. Morrill, and C. Retoré (eds.), Springer-Verlag, 2001.
- [7] C. Casadio and J. Lambek (eds.), *Recent Computational Algebraic Approaches to Morphology and Syntax*, Polimetria, Milan, 2008.
- [8] C. Casadio and M. Sadrzadeh, 'In Search for a Uniform Pattern of Clitic Movement', to appear in the Lecture Notes in Artificial Intelligence (LNAI) proceedings of the Tbilisi Symposium on Logic Language Information, Sept. 2009.
- [9] R. A. Hall Jr., *Hungarian Grammar*, Language Monographs **21**, Linguistic Society of America, 1944.
- [10] J. Lambek, 'Compact Monoidal Categories from Linguistics to Physics', in *New Structures for Physics*, Lecture Notes in Physics, B. Coecke (ed.), Springer, 2009.
- [11] J. Lambek, 'Pregroups and Natural Language Processing', *Springer Science + Business Media* **28**, 2006.
- [12] J. Lambek, 'Iterated Galois Connections in Arithmetic and Linguistic', in *Galois Connections and Applications*, K. Denecke et al. (eds.), 2004.
- [13] J. Lambek, 'Type Grammar Revisited', In A. Lecomte et al. (eds.), *Logical Aspects of Computational Linguistics*, Springer LNAI 1582, 1999.
- [14] B. Megyesi, 'The Hungarian Language, A Short Descriptive Grammar', Department of Linguistics, Stockholm University, 1998, <http://stp.ling.uu.se/~bea/publ/megyesi-hungarian.pdf>.
- [15] M. Sadrzadeh, 'Pregroup Analysis of Persian Sentences', in [7].
- [16] D.N. Yetter, 'Quantales and (non-commutative) Linear Logic', *Journal of Symbolic Logic* **55**, 1990.
- [17] 'Hungarian Grammar', Wikipedia the free encyclopedia, http://en.wikipedia.org/wiki/Hungarian_grammar.

OXFORD UNIVERSITY COMPUTING LABORATORY AND WOLFSON COLLEGE

Current address: Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom.

E-mail address: mehrrs@comlab.ox.ac.uk