

# Clitic Movement in Pregroup Grammar: a Cross-linguistic Approach

Claudia Casadio and Mehrnoosh Sadrzadeh\*

<sup>1</sup>Dept. of Philosophy, Chieti University, IT  
`casadio@unich.it`

<sup>2</sup>Computing Laboratory, Oxford University, UK  
`mehrs@comlab.ox.ac.uk`

**Abstract.** The calculus of pregroups has been introduced by Lambek [10] as an algebraic (and logical) procedure for generating the grammatical analysis of natural languages; it has been applied to a wide range of languages from English and German, to French and Italian, and many others [7]. Pregroups are non-commutative structures, but the syntax of natural languages shows the presence also of cyclic patterns, e.g. those caused by the the so called movements of clitic pronouns. In this paper we propose an extension of the calculus of pregroups including two cyclic meta-rules and use them to formally analyze movement of clitic clusters in Persian, French, and Italian. We also point out that these rules are inspired by Yetter’s and Abramski’s cyclic rules for Linear Logic.

**Keywords.** Type Grammars, Pregroup, Clitic Movement, Cyclic Rules

## 1 Introduction

Pregroups are mathematical structures introduced by Lambek [10, 12, 11] as a procedure for generating the syntax of natural languages alternative to his *Syntactic Calculus* [9]; the latter has been used in linguistics in the context of categorical and type logical grammars, see [14]. The mathematical and logical properties of pregroups have been studied in [4, 5, 10], and they have been applied to the grammatical analysis of a large number of languages: from English, French, Italian, and German to Arabic, Polish, Japanese, Persian and many more, see [7].

The story of this paper is motivated by the following observations: in Persian the subject and object of a sentence occur in pre-verbal position (Persian is a SOV language), but they may attach themselves as clitic pronouns to the end of the verb and form a one-word sentence. By doing so, the word order changes from SOV to VSO. In Arabic we find an analogous pattern, but the word order remains VSO in both cases. A similar phenomenon can be found in some European languages, but in a less extreme format. For example, in Italian and French nouns or pronouns occurring as complements in post-verbal position,

---

\* The authors are sincerely grateful to Jim Lambek for the suggestions made to this paper in course of elaboration. Support by EPSRC (grant EP/F042728/1) is gratefully acknowledged by M. Sadrzadeh; support by MIUR (60% 2009) by C. Casadio.

can take a clitic form and move to a pre- (or post-) verbal position, possibly with a change in their order of appearance. To account for these movements, Casadio and Lambek in [6] and Bagelli and Lambek in [2] had worked with inflectors and extended vs. short infinitive forms of the verb. This resulted in assigning the verbs two different types of infinitive  $j$  and  $i$ , and in defining somewhat lengthy types for clitic pronouns, for example in Italian the type  $(jo^l i^l)$  for the clitic pronoun “lo” (*him, it*) in accusative case. But in so doing, they had managed to keep the type of the verb unchanged and take advantage of the iterated adjoints such as  $o^l$  that are allowed by the pregroup grammar<sup>1</sup>.

In this paper we propose a different approach that simplifies the above analyses for Italian and French and provides a new solution for Persian, offering a unified account. We show how one can explain these movements by adding two *clitic meta-rules* to the pregroup grammars of these languages. In the pregroup extended with these rules the clitic type of the verb is derivable from its original type. Interestingly, the same rule works for (count-noun) clitics in French and pre-verbal clitic clusters in Italian, whereas the converse of the rule works for Persian. In Arabic and post-verbal Italian, in turn, we don’t need these rules, since the clitic forms preserve the order of their non-clitic peers.

Our clitic rules remind one of the *cyclic* rules of Yetter’s Cyclic (Linear) Logic [17]. These rules were later added to Lambek’s Syntactic Calculus by Abrusci [1], proving that a non-commutative calculus with these rules is still a conservative extension of the Syntactic Calculus. Motivated by this fact, we originally wanted to add these rules as extra axioms to a pregroup, or as extra rules to its sequent calculus. Unfortunately, as pointed out by Lambek, the resulting calculus is not conservative anymore: the addition of the cyclic rules makes the pregroup collapse into an ordered group.

## 2 Two Clitic Meta-Rules

A pregroup  $P$  is a partially ordered non-commutative monoid  $(P, \cdot, 1, \leq, (-)^l, (-)^r)$  where each element  $p \in P$  has both a left adjoint  $p^l$  and a right adjoint  $p^r$ , to the extent that the following rewritings hold:

$$p^l p \leq 1 \quad pp^r \leq 1 \quad 1 \leq pp^l \quad 1 \leq p^r p .$$

The first two inequalities are referred to as *contractions* and the second two as *expansions*. One can show that the adjoints are unique and contravariant, i.e.

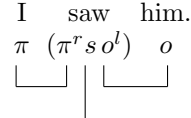
$$p \leq q \implies q^l \leq p^l \quad \text{and} \quad q^r \leq p^r ,$$

the unit 1 and the multiplication (the dot ‘.’ here omitted) are self dual [10, 5]:

$$1^l = 1 = 1^r \quad (pq)^l = q^l p^l \quad (pq)^r = q^r p^r .$$

<sup>1</sup> Pregroups are non-conservative extensions of non-commutative linear logic [1, 10], in which “left” and “right” iterated negations (equivalently “left” and “right” iterated adjoints) do not cancel.

To develop the analysis of a natural language, we define a set of basic types, in particular  $n$  for *noun phrase*,  $\pi$  for *subject*,  $o$  for *object*,  $\omega$  for *indirect object*,  $\lambda$  for *locative object*,  $s$  for *sentence*,  $q$  for *question*, assuming the partial orders  $n \leq \pi, n \leq o, n \leq \omega, n \leq \lambda$ ; we then generate a free pregroup from these basic types and the partial orders between them. The pregroup will have simple types such as  $n^l, n^r, \pi^l, \pi^r, o^l, o^r$ , and compound types such as  $(\pi^r s o^l)$ . For example, the sentence ‘I saw him.’ is typed as follows



A sentence is grammatical iff it *reduces* to the type  $s$  (or  $q$  in the case of questions), a procedure depicted by the under-link diagrams. The monoid operation of the pregroup is non-commutative, but if we add the following meta-rules, then we obtain a limited form of commutativity, for  $p, q \in P$

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

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

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 stressed pronouns or arguments. For any clitic pronoun  $p$ , 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  $\bar{p}q = \bar{p} \bar{q}$ . The mention of the word *original* in the rules is to forbid the application of the rule to the derived types of the verb. We want to be able to apply the rule only to the type of the verb in the full sentence without any clitic pronouns. If we do not assume this restriction and allow the rule to be applied to the type of the verb also when clitics occur, then non-grammatical sentences can be formed. We will give examples in the following sections. Note that these rules are converse to each other, not considering the use of the over-lined types.

### 3 Clitic Movement in Persian

The clitic clusters (pre-verbal vs. post-verbal) for the sentence *I saw him*, “man u-ra didam” in Persian, exhibit the following general pattern:



The over-lined types  $\bar{\pi}, \bar{o}$ , stand for the clitic versions of the subject and object pronouns. The “am” at the end of “didam” in the non-clitic sentence above is a

conjugation pronoun and liaises with the first person subject “man”. Conjugation pronouns are also used as clitic pronouns for the subject. Typing the conjugation and the accord can be done by indexing the subject and the  $s$  in the verb with its person and require that they should cancel out only if they have the same index. Making the typing for conjugation pronouns explicit would be distracting for the purpose of this paper, so we shall ignore it and focus on the clitic pronouns. Then the conjugation pronoun “am” in the first sentence is included into the type of the verb, whereas the “am” in the second sentence, being a clitic pronoun, receives a type.

We postulate the first clitic rule as a meta-rule in the pregroup grammar of Persian. Applying this rule to the above example will allow us to obtain the clitic form of the verb from its original type. The original Persian verb has the type  $o^r \pi^r s = (\pi o)^r s$ , which is of the form  $p^r q$ ; after applying the clitic rule (1) we obtain  $s(\pi \bar{o})^l = s(\pi \bar{o})^l = s \bar{o}^l \bar{\pi}^l$ , i.e. the type of the verb with postverbal clitics<sup>2</sup>. The clitic rule can be seen as a re-write rule and the derivation can be depicted as one-linear as follows

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

To form these one-word sentences, one needs not to have pronouns as subject and object in the original sentence. They can as well be nouns or noun phrases, for example the sentence *I saw Nadia*, in Persian “man Nadia-ra didam”, becomes “did-am-ash”, being typed exactly as in the example above. This movement takes a more subtle form when the subject is third person singular, since in this case the subject clitic pronoun is the same as the last letter of the verb. But the types and the movement are exactly the same as in our previous examples.

Hassan	Nadia	saw		saw	he	her
Hassan	Nadia-ra	did.		di	d	ash.
$\pi$	$o$	$(o^r \pi^r s) \rightarrow s$		$(s \bar{o}^l \bar{\pi}^l)$	$\bar{\pi}$	$\bar{o} \rightarrow s$

One can form a yes-no question from any of the sentences above, by adding the question form “aya” to the beginning of the sentence. Since in Persian the word order of the question form is the same as that of the original sentence, the clitic movement remains the same and obeys the same rule, as follows

Did	Hassan	Nadia	see?	Did	see	he	her?
aya	Hassan	Nadia-ra	did?	aya	di	d	ash?
$qs^l$	$\pi$	$o$	$(o^r \pi^r s) \rightarrow q$	$qs^l$	$(s \bar{o}^l \bar{\pi}^l)$	$\bar{\pi}$	$\bar{o} \rightarrow q$

### 3.1 Partial Clitic Movement

One can also obtain partial clitic movement, like in *Hassan saw her*, in Persian “Hassan did-ash”. This would make sense as an answer to the wh-question *Who saw Nadia?* that in Persian translates as “Che-kasi u-ra did?”

<sup>2</sup> In the pregroup calculus, like in non-commutative linear logic, a form of De Morgan rule holds to the extent that  $(pq)^l = (q^l p^l)$  and  $(pq)^r = (q^r p^r)$ ; see [1, 10].

Hassan	saw	her.
Hassan	did	ash.
$\pi$	$(\pi^r s \bar{o}^l)$	$\bar{o} \rightarrow s$

This partial clitic type of the verb is derivable via the same clitic rule, but applied in a different manner. Here one takes  $p^r$  to be  $o^r$  and  $q$  to be  $\pi^r s$ , the application of the rule results in  $q\bar{p}^l$  equal to  $(\pi^r s) \bar{o}^l = (\pi^r s \bar{o}^l)$ , via associativity of ‘.’.

Another example of a partial clitic movement is the sentence “Nadia saw I” (*I saw Nadia*), in Persian “Nadia-ra did-am”. This is more controversial than the previous example, and might not be considered as grammatical, as it is a phrase that starts neither with a subject nor with a verb. Perhaps it only makes sense as an immediate answer to the question *Whom did I see?*, in Persian “Man che-kasi-ra did-am?”.

Nadia	saw	I.
Nadia-ra	did	am.
$o$	$(o^r s \bar{\pi}^l)$	$\bar{\pi} \rightarrow s$

Note that this new type  $(o^r s \bar{\pi}^l)$  of the verb is not derivable from its original type  $(o^r \pi^r s)$  and that the sentence is not fully grammatical; one should rather say “man Nadia ra did-am”. However, if one insists on a solution, one needs to generalize our clitic rule to the effect that it can be applicable to parts of the type string of the verb. Here is the definition of such a *partial clitic rule*:

For  $p, q \in P$ , if  $p^r q$  is in the *original* type of the verb, then so is  $q\bar{p}^l$ .

This means that if a string of the type  $p^r q$  occurs somewhere inside the type of a verb, it can be replaced with a string of the type  $q\bar{p}^l$  and the resulting type is considered as another type for that verb.

One can now apply this rule to the  $\pi^r s$  part of the full type of the verb, that is  $(o^r \pi^r s)$ , and obtain  $(s \bar{\pi}^l)$ , leaving the  $o^r$  at the beginning untouched. In this way we obtain the type that we need, that is  $(o^r s \bar{\pi}^l)$ . Despite its scary appearance, the partial clitic rule is harmless. The only other way one can apply it to the original type of the verb is by fixing  $s$  at the end and taking  $p^r = o^r$  and  $q = \pi^r$ , applying the rule will then result in  $(\pi^r \bar{o}^l s)$ , which cannot make any ungrammatical sentence grammatical. This is due to the occurrence of the left adjoint  $\bar{o}^l$  of the clitic pronoun, which needs a double left adjoint  $\bar{o}^{ll}$  to cancel out, but we have assigned no grammatical role to types such as  $\bar{o}^{ll}$  in this paper.

One can use this type to form yes-no questions from the above question *Did I See Nadia?*, in Persian “aya Nadia-ra did-am?”, etc.

Did	Nadia	see	I?
aya	Nadia-ra	did	am?
$qs^l$	$o$	$(o^r s \bar{\pi}^l)$	$\bar{\pi} \rightarrow q$

The grammatical form of this question would be “aya man Nadia ra did-am?”, just like the grammatical form of the original sentence.

Had we allowed for the application of the clitic rule to the derived types of the verb, rather than only to its original type, we would have been able to apply

it to the type of the verb in the partial clitic sentence above for  $p^r q = o^r(s\bar{\pi}^l)$ , by taking  $p^r = o^r$  and  $q = s\bar{\pi}^l$  and obtaining  $q\bar{p}^l = s\bar{\pi}^l\bar{o}^l$ , which would have made the ungrammatical one-word sentence “did-ash-am” grammatical.

saw	him	I
*did	ash	am
$s\bar{\pi}^l\bar{o}^l$	$\bar{o}$	$\bar{\pi}$

### 3.2 Locative Objects and Auxiliaries

One can introduce a locative object in the example above and utter *I saw him in the library*, in Persian “man u-ra dar-ketabkhaneh didam”. There are no clitic pronouns for the locative object in Persian, so the *short-form* of this sentence is not one word any more. One can still move the pronouns and form shorter sentences, but in this case the locative object “dar-ketabkhaneh” stays before the verb and only the subject and object pronouns become post-verbal.

I	him	in the library	saw	in the library	saw	I	him
man u-ra	dar-ketabkhaneh	didam.		dar-ketabkhaneh	did	am	ash.
$\pi$	$o$	$\lambda$	$(\lambda^r o^r \pi^r s) \rightarrow s$	$\lambda$	$(\lambda^r s \bar{o}^l \bar{\pi}^l)$	$\bar{\pi}$	$\bar{o} \rightarrow s$

Similarly to the second example of partial clitic movement, this new type for the verb is not derivable from its original type by means of the clitic rule, since according to this rule, one can only move the types from the beginning or the end of the type string. There is no way to pull out the type  $s$  from the middle of  $(\lambda^r s \bar{o}^l \bar{\pi}^l)$  and move it to the end of the type string in  $(\lambda^r o^r \pi^r s)$ , while keeping the initial type  $\lambda^r$  untouched. If one insists on applying the rule, from  $(\lambda^r o^r \pi^r s)$  one obtains  $(s\bar{\lambda}^l \bar{o}^l \bar{\pi}^l)$ , which would only type check if there was a clitic pronoun for location in Persian, something similar to the  $y$  in French. But this is not the case in Persian. This may be again good news, because while there is no doubt that one word sentences like “did-am-ash” are grammatical, the two word ones “dar-ketabkhaneh did-am-ash” are controversial; in fact, it would be a hard job to find a Persian grammar book which claims that the latter is grammatical. Its use is perhaps only justified as the short answer to the question *Where did you see him?*, in Persian “dar-koja did-i-ash?”, which is itself a short form of the full question “To u-ra dar-koja didi?”. If one does not agree with the above argument, one can use the partial version of the clitic rule, the **in** rule given above: we apply it only to the  $(o^r \pi^r s)$  part of the verb and obtain  $(s\bar{o}^l \bar{\pi}^l)$ , without touching the type  $\lambda^r$  at the beginning.

Persian sentences in perfect tenses are built by using auxiliary verbs and either the participle or the subjunctive form of the verb. A full analysis of the Persian tenses has been provided in [8]; we skip the details here and, for the sake of simplicity and in order to focus on clitic movement, we introduce the types  $p$  for the *past participle* and  $j$  for the *subjunctive* forms of the verb. Sentences in perfect tense also allow for clitic movement, for example the following sentences in the *past perfect*. The second one is obtained from the first one via clitic

movement: the clitic pronouns attach themselves to the end of the auxiliary verb “budam” (*had*).

I	him	seen	had		seen	had	I	him
man	u-ra	dideh	budam.		dideh	bud	am	ash.
$\pi$	$o$	$p$	$(p^r o^r \pi^r s) \rightarrow s$		$p$	$(p^r s \bar{o}^l \bar{\pi}^l)$	$\bar{\pi}$	$\bar{o} \rightarrow s$

The type of the verb of the second sentence is obtained from the verb type of the first one by applying the partial **in** rule: fix  $p$  and move the  $(o^r \pi^r) = (\pi o)^r$  from the left of  $s$  to its right, so it becomes  $(\bar{\pi} \bar{o})^l = \bar{o}^l \bar{\pi}^l$ . As another example, consider the following sentence in present subjunctive form, involving the modal verb “mikhaham” (*want*) and the verb “bebinam” (*I see*) to express a wish. We treat the subjunctive form of the verb similarly to its infinitive. Like in French, this form needs an object but no external subject; it is the modal verb instead that takes the external subject and a subjunctive  $j$  to produce the sentence  $s$ .

I	want	him	see		want	I	see	him
man	mikhaham	u-ra	bebinam.		mikhah	am	bebinam	ash.
$\pi$	$(\pi^r s j^l)$	$o$	$(o^r j) \rightarrow s$		$(s j^l \bar{\pi}^l)$	$\bar{\pi}$	$(j \bar{o}^l)$	$\bar{o} \rightarrow s$

Clitic movement takes an interesting form here: since the modal verb inputs the subject, clitic movement ends up into the form “mikhah-am”:  $(\pi^r s j^l)$  becomes  $(s j^l \bar{\pi}^l)$ ; and since the subjunctive form of the verb inputs the object, objective clitic movement results in “bebinam-ash”:  $(o^r j)$  becomes  $(j \bar{o}^l)$ .

### 3.3 An Insight into Cliticization in Arabic

We conclude this section by taking a brief look at Arabic. In general (not considering word order changes caused by putting into focus different parts of speech) Arabic is a VSO language. Consider the following transitive sentence

carries	Hassan	a book.
yahmela	Hassan	ketaban.
$(s o^l \pi^l)$	$\pi$	$o \rightarrow s$

where “yahmela” is the verb, “Hassan” is the subject, and “ketaban” is the object. In its short form, this sentence becomes just one word: “yahmel-a-ha”, but the order remains the same VSO, see [15, 3].

carries	he	it.
yahmel	a	ha.
$(s \bar{o}^l \bar{\pi}^l)$	$\bar{\pi}$	$\bar{o} \rightarrow s$

Here the vowel “a” stands for the clitic form of the subject pronoun and refers to “Hassan”, “ha” is the clitic form of the object pronoun and refers to “ketaban”. Since there is no movement, we need not use the clitic meta-rules. The clitic type of the verb is easily derivable from its original type by applying the postulate  $\bar{x} \leq x$ , the conventions about bars  $\bar{p} \bar{q} = \overline{p q}$ , and by remembering that adjoints reverse the order  $x^l \leq \bar{x}^l$ . The derivation is as follows

$$s o^l \pi^l = s(\pi o)^l \leq s(\bar{\pi} \bar{o})^l = s(\bar{\pi} \bar{o})^l = s \bar{o}^l \bar{\pi}^l$$

## 4 Clitic Movement in French

In French, clitic clusters move in the opposite direction with respect to Persian. We will therefore invoke the second clitic meta-rule on the basis of which we derive the type of the clitic form of the verb from its original type. Let's start with a simple example, the sentence “Jean voit Marie” (*Jean sees Marie*) and its clitic form “Jean la voit”. Following the Persian examples, we derive the type  $(\bar{o}^r \pi^r s)$  of the clitic form of the verb starting from its original type  $(\pi^r s o^l)$ : we take  $q = (\pi^r s)$ ,  $p^l = o^l$ , apply the clitic rule (2) and obtain  $(\pi^r s o^l) = (\pi^r s) o^l \rightsquigarrow \bar{o}^r (\pi^r s) = (\bar{o}^r \pi^r s)$ . Here are the derivations for the two sentences

$$\begin{array}{ll} \text{Jean} & \text{voit} & \text{Marie.} & & \text{Jean} & \text{la} & \text{voit.} \\ \pi & (\pi^r s o^l) & o & \rightarrow s & \pi & \bar{o} & (\bar{o}^r \pi^r s) \rightarrow s \end{array}$$

Consider a second simple example, this time with the locative object  $\lambda$  and its clitic pronoun  $\bar{\lambda}$ .

$$\begin{array}{ll} \text{Jean} & \text{va} & \text{à Paris.} & & \text{Jean} & \text{y} & \text{va.} \\ \pi & (\pi^r s \lambda^l) & \lambda & \rightarrow s & \pi & \bar{\lambda} & (\bar{\lambda}^r \pi^r s) \rightarrow s \end{array}$$

Again the clitic rule easily derives  $(\bar{\lambda}^r \pi^r s)$  from  $(\pi^r s \lambda^l)$ . Now consider the more complicated example “Jean donne une pomme à Marie.” (*Jean gives an apple to Marie*), we type it as follows

$$\begin{array}{ll} \text{Jean} & \text{donne} & \text{une pomme} & \text{à Marie.} \\ \pi & (\pi^r s w^l o^l) & o & w \rightarrow s \end{array}$$

While learning French at school, we always had difficulty in remembering the order of the clitic pronouns in this kind of sentences. Our clitic rule (2) offers a solution: according to it a verb of type  $(\pi^r s w^l o^l)$  can also be of type  $(\bar{w}^r \bar{o}^r \pi^r s)$ , taking  $q = (\pi^r s)$ ,  $p = (ow)^l$  and deriving  $(\pi^r s w^l o^l) = (\pi^r s)(w^l o^l) = (\pi^r s)(ow)^l \rightsquigarrow (\bar{o}w)^r (\pi^r s) = (\bar{w}^r \bar{o}^r)(\pi^r s) = (\bar{w}^r \bar{o}^r \pi^r s)$ . This type for the clitic verb form allows one to obtain grammatical sentences like the following

$$\begin{array}{ll} \text{Jean} & \text{la} & \text{lui} & \text{donne.} \\ \pi & \bar{o} & \bar{w} & (\bar{w}^r \bar{o}^r \pi^r s) \rightarrow s, \end{array}$$

but it will *not* allow the following incorrect order

$$\begin{array}{ll} \text{Jean} & \text{lui} & \text{la} & \text{donne} \\ \pi & \bar{w} & \bar{o} & (\bar{w}^r \bar{o}^r \pi^r s) \end{array}$$

We can also predict clitic movement within the sentence “Jean met une pomme sur la table” (*Jean puts an apple on the table*), assigning the type  $(\pi^r s \lambda^l o^l)$  to the original verb. Clitic rule (2) assures that this verb can also receive the type  $(\bar{\lambda}^r \bar{o}^r \pi^r s)$ , allowing the derivation of grammatical sentences like the following

$$\begin{array}{ll} \text{Jean} & \text{la} & \text{y} & \text{met.} \\ \pi & \bar{o} & \bar{\lambda} & (\bar{\lambda}^r \bar{o}^r \pi^r s) \rightarrow s \end{array}$$

but *not* of the following ungrammatical one



$$\begin{array}{ccccccc} \text{Jean} & \text{y} & \text{la} & & \text{met} & & \\ \pi & \bar{\lambda} & \bar{o} & (\bar{\lambda}^r \bar{o}^r \pi^r s) & & & \end{array}$$

As a more complicated example, consider a sentence with a modal verb like “Jean peut donner une pomme à Marie” (*Jean can give an apple to Marie*), which we type as follows

$$\begin{array}{ccccccc} \text{Jean} & \text{peut} & \text{donner} & \text{une} & \text{pomme} & \text{à} & \text{Marie.} \\ \pi & (\pi^r s i^l) & (i w^l o^l) & & o & & w \rightarrow s \end{array}$$

This time we apply the clitic rule to the infinitive of the verb  $(i w^l o^l)$  and obtain  $(\bar{w}^r \bar{o}^r i)$  for its clitic form. This types the following sentence

$$\begin{array}{ccccccc} \text{Jean} & \text{peut} & \text{la} & \text{lui} & \text{donner.} & & \\ \pi & (\pi^r s i^l) & \bar{o} & \bar{w} & (\bar{w}^r \bar{o}^r i) & \rightarrow s & \end{array}$$

But *not* the following incorrect one

$$\begin{array}{ccccccc} \text{Jean} & \text{peut} & \text{lui} & \text{la} & \text{donner} & & \\ \pi & (\pi^r s i^l) & \bar{w} & \bar{o} & (\bar{w}^r \bar{o}^r i) & & \end{array}$$

The same technique of applying the rule to the type of the infinitive of the verb works for simpler sentences such as “Jean peut voire Marie”.

In order to type clitic movement for sentences in the past tense like “Jean a donné une pomme à Marie” (*Jean gave an apple to Marie*), we apply the clitic rule to the verb phrase that is the combination of the auxiliary and the past-participle of the verb (see [2, 6]). Having typed the sentence as follows

$$\begin{array}{ccccccc} \text{Jean} & \text{a} & \text{donné} & \text{une} & \text{pomme} & \text{à} & \text{Marie.} \\ \pi & (\pi^r s p^l) & (p w^l o^l) & & o & & w \rightarrow s \end{array}$$

we need to take the type of “a donné”, which is  $(\pi^r s p^l)(p w^l o^l) \leq (\pi^r s w^l o^l)$ , after the contraction  $p^l p \leq 1$ , then we take  $q = (\pi^r s)$ ,  $p^l = (o w)^l$ , and apply the clitic rule (2) to obtain  $\bar{p}^r q = \bar{w}^r \bar{o}^r \pi^r s$ . The clitic sentence is thus typed as follows

$$\begin{array}{ccccccc} \text{Jean} & \text{la} & \text{lui} & \text{a} & \text{donnée.} & & \\ \pi & \bar{o} & \bar{w} & (\bar{w}^r \bar{o}^r \pi^r s) & \rightarrow s & & \end{array}$$

#### 4.1 Partial Clitic Movement

In the above, we have analyzed sentences where both objects are replaced with clitic pronouns. This need not always be the case. For instance our example sentence “Jean donne une pomme à Marie”, can also turn into the sentence “Jean la donne à Marie” by partial clitic movement. The type  $(\bar{o}^r \pi^r s w^l)$  of the verb in the latter sentence is obtainable from the original type of the verb  $(\pi^r s w^l o^l)$  by applying clitic rule (2) as before, for  $q = (\pi^r s w^l)$  and  $p^l = o^l$ . However the same will not work for the other partial clitic sentence “Jean lui donne une pomme”, where we need the type  $(\bar{w}^r \pi^r s o^l)$  for the verb. Similarly to the Persian case, to derive this type, we generalize clitic rule (2) as follows

For  $p, q \in P$ , if  $qp^l$  is in the *original* type of the verb then so is  $\bar{p}^r q$ .

With this version of the rule one can derive  $(\bar{w}^r \pi^r s o^l)$  from  $(\pi^r s w^l o^l)$ , by applying the rule to the  $(\pi^r s w^l)$  part of the type string. To do so, one takes  $q = (\pi^r s)$  and  $p^l = w^l$ ; after applying the in rule one obtains  $(\bar{w}^r \pi^r s)$ , to which one then glues the type  $o^l$  at the end of the type string.

A similar problem occurs with locative clitic movement in “Jean met une pomme sur la table”. The partial clitic movement giving “Jean la met sur la table” works as before: the new type of the verb  $(\bar{o}^r \pi^r s \lambda^l)$  is derivable from its original type  $(\pi^r s \lambda^l o^l)$ . But for the other partial movement “Jean y met une pomme”, one needs to use the in rule to obtain  $(\bar{\lambda}^r \pi^r s o^l)$ : we fix the  $o^l$  at the end of the type string, then apply the rule to the  $(\pi^r s \lambda^l)$  part deriving  $(\bar{\lambda}^r \pi^r s)$ .

Had we allowed the clitic rule to apply to non-original types of the verb, then we would have been able to apply the rule also to its derived types, for example those assigned to a sentence with a partial clitic movement. In the above examples these types would be  $(\bar{w}^r \pi^r s o^l)$  or  $(\bar{\lambda}^r \pi^r s o^l)$ . The application of the rule to these types would have derived the types  $(\bar{o}^r \bar{w}^r \pi^r s)$  and  $(\bar{o}^r \bar{\lambda}^r \pi^r s)$ , which would have respectively made the ungrammatical sentences “Jean lui la donne” and “Jean y la met” grammatical<sup>3</sup>.

## 5 Clitic Movement in Italian

Clitic movement in Italian exhibits two general patterns: clitic pronouns can occur both in pre-verbal and post-verbal position, keeping the same relative order: locative/indirect object, direct object. Differently from Persian, subjects do not allow clitic counterparts and concerning the other verbal arguments, clitic unstressed elements can be attached both to the main verb or to auxiliaries and modal verbs. Therefore clitic types will include  $\lambda$ , the type of the locative object,  $\omega$  the type of the indirect object (dative), and  $o$  the type of the direct object (accusative).

Sentences with one occurrence of a pre-verbal clitic can be obtained exactly like in French, as shown in the following examples corresponding to the French sentences given above “Gianni vede Maria” and its clitic form “Gianni la vede”.

$$\begin{array}{ccc} \text{Gianni} & \text{vede} & \text{Maria.} \\ \pi & (\pi^r s o^l) & o \rightarrow s \end{array} \qquad \begin{array}{ccc} \text{Gianni} & \text{la vede.} \\ \pi & \bar{o} & (\bar{o}^r \pi^r s) \rightarrow s \end{array}$$

To derive the clitic type of the verb we start with the original type  $(\pi^r s o^l)$ , take  $q = (\pi^r s)$ ,  $p^l = o^l$ , apply the clitic rule (2) and obtain the type  $(\bar{o}^r \pi^r s)$ . The

<sup>3</sup> The patterns that does not follow from the above rules for clitic movement in French is when the direct object is a mass noun, e.g. in “Jean donne de la farine à Marie” with clitic form “Jean lui en donne”. This pattern also occurs in Italian, so we refer to the next section for some comments; see also [2, 13]. One might think that sentences when the indirect object is not a noun, e.g. “Jean donne une pomme à nous”, would also be exceptions, as their clitic form would be “Jean nous la donne”, but fortunately these sentences are not fully grammatical to start with.

same process applies with a locative argument  $\lambda$  and the corresponding clitic pronoun  $\bar{\lambda}$ , where the clitic rule derives  $(\bar{\lambda}^r \pi^r s)$  from  $(\pi^r s \lambda^l)$ .

$$\begin{array}{ccc} \text{Gianni} & \text{va} & \text{a Roma.} \\ \pi & (\pi^r s \lambda^l) & \lambda \rightarrow s \end{array} \qquad \begin{array}{ccc} \text{Gianni ci} & \text{va.} \\ \pi & \bar{\lambda} & (\bar{\lambda}^r \pi^r s) \rightarrow s \end{array}$$

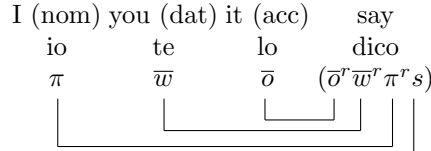
When we consider more complicated cases of a verb with two arguments like in “Gianni da un libro a Maria” (*Gianni gives a book to Maria*), or “Gianni mette un libro sul tavolo” (*Gianni puts a book on the table*) we find that the clitics pronouns occur in the opposite order with respect to French<sup>4</sup>: e.g. the verb “dare” (*to give*) has the clitic form “Gianni glie lo da”, and the verb “mettere” (*to put*) the clitic form “Gianni ce lo mette”.

In [6] this problem was handled by introducing a second type for verbs with two complements in which the arguments order is reversed  $(\pi^r s o^l w^l)$  and  $(\pi^r s o^l \lambda^l)$ <sup>5</sup>; we will take the same move here and show that from these verb types, that we would like to name *prototypes*, all the pre-verbal and post-verbal clitic patterns follow. The advantage of the present analysis is that we will make use only of the clitic metarules and of some order postulates.

Applying clitic rule (2) to the new types we obtain the correct clitic verb forms to handle the cases of pre-verbal cliticization mentioned above:  $(\pi^r s o^l w^l) = (\pi^r s)(wo)^l \rightsquigarrow (\bar{w}\bar{o}^r)(\pi^r s) = (\bar{o}^r \bar{w}^r \pi^r s)$ , and the same with  $\lambda$  in place of  $w$ .

$$\begin{array}{ccc} \text{Gianni glie lo} & \text{da.} \\ \pi & \bar{w} & \bar{o} (\bar{o}^r \bar{w}^r \pi^r s) \rightarrow s \end{array} \qquad \begin{array}{ccc} \text{Gianni ce lo mette.} \\ \pi & \bar{\lambda} & \bar{o} (\bar{o}^r \bar{\lambda}^r \pi^r s) \rightarrow s \end{array}$$

The following diagram shows the general pattern of preverbal cliticization in Italian with a verb taking two arguments:



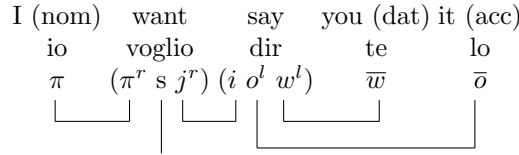
To get the post-verbal clitic forms we do not need apply the clitic meta-rules; in the case of transitive verbs such as “vedere” (*to see*), or motion verbs such as “andare” (*to go*) the post-verbal clitic forms are obtained by the original types  $(io^l)$ ,  $(i\lambda^l)$  giving the order conditions on the pregroup:  $\bar{o} \leq o$ ,  $\bar{\lambda} \leq \lambda$ , on the basis of which an argument of type  $\bar{o}$  is also of type  $o$ , and an argument of type  $\bar{\lambda}$  is also of type  $\lambda$  [10, 6].

<sup>4</sup> In French the order of the accusative and dative/locative third person clitics is reversed with respect to Italian and other Romance languages such as Spanish or Portuguese; but the order is similar in the case of first or second person pronouns (see footnote 4).

<sup>5</sup> These types are admitted by other languages such as English where you find sentences such as “John gives Mary a kiss” or “John took to the station the bag”, and we can consider them as reminiscent of the basic SOV order holding for Latin, or other Indo-European languages.

$$\begin{array}{ll}
\text{Gianni vuole veder}.\text{la} . & \text{Gianni vuole andar}.\text{ci} . \\
\pi \quad (\pi^r s i^l) \quad (i o^l) \quad \bar{o} \rightarrow s & \pi \quad (\pi^r s i^l) \quad (i \lambda^l) \quad \bar{\lambda} \rightarrow s
\end{array}$$

Similarly the post-verbal clitic forms with two arguments follow from the *prototypes* introduced above, by means the same ordering conditions: e.g. assigning the type  $(\pi^r s o^l w^l)$  to the verb “dare” (*to give*) it applies to the clitic types  $\bar{o}$ ,  $\bar{w}$ , under the order condition  $\bar{o} \leq o$ ,  $\bar{w} \leq w$ ; the same with the verb type  $(\pi^r s o^l \lambda^l)$  and the condition  $\bar{\lambda} \leq \lambda$ . The following diagram gives a general picture of the post-verbal construction with a verb taking two arguments, cf. [6].



### 5.1 Partial clitic movement

We conclude this section by showing how partial clitic movement in Italian is obtained within the present approach. Italian allows partial clitic movement in sentences with two arguments verbs like “dare” (*to give*) considered above. Such clitic verb forms are characterized by the presence of a full argument and a clitic pronoun that can be, in turn, a direct or indirect object or a locative. Moreover, the clitic pronoun can occur both in pre-verbal and post-verbal position. The preverbal cases are like in the following examples where the clitic pronoun is the indirect object in the first case and the direct object in the second

$$\begin{array}{ll}
\text{Gianni gli} \quad \text{da} \quad \text{un libro.} & \text{Gianni lo} \quad \text{da} \quad \text{a Maria.} \\
\pi \quad \bar{w} \quad (\bar{w}^r \pi^r s o^l) \quad o \rightarrow s & \pi \quad \bar{o} \quad (\bar{o}^r \pi^r s w^l) \quad w \rightarrow s
\end{array}$$

To derive the first sentence we choose the type  $(\pi^r s o^l w^l)$  for the verb “dare” and take  $p^r = w^l$ ,  $q = (\pi^r s o^l)$ , then by applying clitic rule (2) we obtain  $p^r q = (\bar{w}^r)(\pi^r s o^l) = (\bar{w}^r \pi^r s o^l)$ . For the second case of pre-verbal cliticization we need to invoke the **in** rule given above for French, the content of which is: if a type string  $qp^l$  occurs somewhere inside the *original* type of a verb, it can be replaced by a type string  $\bar{p}^r q$  and the resulting type is considered as another type for that verb. In this case the string in question is  $(\pi^r s o^l)$  where  $q$  is  $(\pi^r s)$  and  $p^l$  is  $o^l$ ; the resulting type string is  $(\bar{o}^l \pi^r s) w^l = (\bar{o}^l \pi^r s w^l)$  that allows the preverbal clitic to apply on the left side of the verb and the indirect argument to apply to the other side. Finally, post-verbal partial clitic movement is so exemplified

$$\begin{array}{ll}
\text{Gianni vuole} \quad \text{dar}.\text{gli} \quad \text{un libro.} & \text{Gianni vuole} \quad \text{dar}.\text{lo} \quad \text{a Maria.} \\
\pi \quad (\pi^r s i^l) \quad (i o^l w^l) \quad \bar{w} \quad o \rightarrow s & \pi \quad (\pi^r s i^l) \quad (i w^l o^l) \quad \bar{o} \quad w \rightarrow s
\end{array}$$

The first sentence with post-verbal partial cliticization follows directly from the verb type and the order condition  $\bar{w} \leq w$ : we can interpret this result as meaning that a post-verbal clitic stays *in situ* like any verb argument. The second sentence

is more subtle, in fact we cannot derive it from the *prototype*. However, we believe that this form of clitic movement has appeared later in the development of modern Italian. Thus it would make more sense to derive it from the modern type of the Italian ditransitive verb, that is  $(\pi^r s w^l o^l)$ . Observe that if we choose this type, we easily derive the type of “vuole dar”, i.e.  $(\pi^r s w^l \bar{o}^l)$  by the  $(i^l i)$  contraction, and the sentence type via the order postulate  $\bar{o} \leq o$ .

## 6 Clitic Rules and Cyclic Pregroups

In the above, we have formulated the clitic rules in a form that Lambek would call *meta-rules*. At some risk of over generation, one is tempted to formulate these rules as axioms and add them to the pregroup calculus (or add their rule version to the sequent calculus of pregroups). These axioms belong to the *cyclic* calculus proposed by Abrusci in [1] in the following way

$$\frac{\vdash \Gamma, \Delta}{\vdash \Delta^{+2}, \Gamma} (rr) \qquad \frac{\vdash \Gamma, \Delta}{\vdash \Delta, \Gamma^{-2}} (ll)$$

Via the standard translation between Syntactic Calculus and Pregroups [10] (positive formulae as right adjoints and negative formulae as left adjoints), the axiomatic version of these rules becomes<sup>6</sup>

$$(1) \quad qp \leq pq^l \qquad (2) \quad qp \leq p^{rr} q$$

We refer to these as *cyclic axioms*, in particular to the first one as the *left cyclic axiom* and to the second one as the *right cyclic axiom*. We can then re-formulate our clitic *meta-rules* as *clitic axioms*

$$\text{Persian} \quad p^r q \leq q\bar{p}^l \qquad \text{French-Italian} \quad qp^l \leq \bar{p}^r q$$

where the latter is derivable from the former, and prove the following results:

**Proposition 1.** *The clitic axioms are derivable from the cyclic axioms.*

**Proof.** The axiom for French is derivable from the *right cyclic axiom* as follows, take  $p$  to be  $p^l$  and observe that  $(p^l)^{rr} = p^r$ , then one obtains  $qp^l \leq p^r q$ . Since  $\bar{p} \leq p$ , and since adjoints are contravariant, we have  $p^r \leq \bar{p}^r$ , thus  $p^r q \leq \bar{p}^r q$ , and by transitivity of order we obtain  $qp^l \leq \bar{p}^r q$ . The axiom for Persian and Italian is derivable from the *left cyclic axiom* as follows: take  $q$  to be  $p^r$  and  $p$  to be  $q$ . Now since  $(p^r)^l = p^l$ , we obtain  $p^r q \leq qp^l$ , and since  $\bar{p} \leq p$ , being adjoints contravariant, we have  $p^l \leq \bar{p}^l$ , thus  $qp^l \leq q\bar{p}^l$ , and by transitivity of order we obtain  $p^r q \leq q\bar{p}^l$ .

It is interesting that the rules for clitic movement correspond to well known logical rules of cyclicity. Accordingly, one may call French and Italian *right cyclic* languages and Persian a *left cyclic* language. The consequences of enriching a pregroup with these cyclic axioms (or rules) are however not so desirable.

<sup>6</sup> One can also proceed with the rules rather than their axiomatic version.

**Proposition 2.** *(due to J. Lambek) A pregroup  $P$  with either of the cyclic axioms is a partially ordered group.*

**Proof.** Consider the left cyclic axiom; if one takes  $q = 1$ , we obtain  $p^r \leq p^l$  for all  $p \in P$ , from which one obtains  $p^{ll} \leq p$ . 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  $w^r = w^l$ . A similar argument can be made for the right cyclic axiom.

Although as proven by Abrusci, Cyclic Bilinear Logic is a conservative extension of Bilinear logic, this is not the case for Cyclic Compact Bilinear Logic and Compact Bilinear Logic (the logical calculus of Pregroups) [5, 4, 12].

## 7 Conclusions

We have tried to find a general pattern for clitic movement in Persian, Italian, and French. We have developed two general rules for this movement and postulated them as meta-rules in the pregroup grammar of each language. We have applied these rules to several examples and also checked that they do not produce ungrammatical strings. It seems to us that a general pattern of clitic movement exists for these languages: French and Italian are *right-cyclic* languages whereas Persian is *left cyclic*. That is, the clitics in French and Italian move in the opposite direction with respect to those in Persian. We faced the challenge of parsing sentences with partial clitic movement in all of these languages, and good news is that these challenges were overcome successfully, again using a uniform pattern of partial applications of the rules. These again are exactly the same in French and Italian, and apply in the opposite direction in Persian.

We have shown why one should resist temptation and introduce these rules as *meta-rules* that can only apply to specific type strings, namely those associated with (the finite or infinitive forms of the) verbs, that we call *original* types. In other words, one should not add these rules as extra axioms to the pregroup or as extra rules to its calculus, in the spirit of the cyclic axioms of Yetter [17] or the cyclic rules of Abrusci [1]. This is because these additions would force the pregroup to reduce to an ordered group.

Let us conclude by reconciling our analysis with previous work on French [2] and Italian [6]. There *iterated* adjoints were used to type the clitic pronouns matching their movements in a sentence. However, we can see that iterated adjoints show up in our work too, since as observed by Lambek, the  $\bar{p}^r$  used in the meta-rule for French and Italian is nothing but  $(\bar{p}^l)^{rr}$ , and the  $\bar{p}^l$  used for Persian is nothing but  $(\bar{p}^r)^{ll}$ . Another way to recover the iterated adjoints in a more explicit way would be to assign the type  $x^{ll}$  to the clitic pronouns for  $x \in \{o, w, \lambda\}$  (instead of  $\bar{x}$ ) and re-formulate the clitic meta-rules and make  $p^r q$  substitutable with  $qp^r$  in Persian and  $qp^l$  with  $p^l q$  in French and Italian. We leave a comprehensive investigation in this direction to future work.

## References

1. Abrusci, M.: Classical Conservative Extensions of Lambek Calculus. *Studia Logica*, 71, 277–314 (2002)
2. Bargelli, D., Lambek J.: An Algebraic Approach to French Sentence Structure. In De Groote, P., Morrill, G., Retoré, C. (eds.) *Logical Aspects of Computational Linguistics*. LACL 2009, pp. 62–78. Springer, Berlin (2001)
3. Bargelli, D., Lambek, J.: An Algebraic Approach to Arabic Sentence Structure. *Linguistic Analysis*, 31 (2001)
4. Barr, M.: On Subgroups of The Lambek Pregroup. *Theory and Application of Categories* 12(8), 262–269 (2004)
5. Buszkowski, W.: Lambek Grammars Based on Pregroups. In: De Groote, P., Morrill, G., Retoré, C. (eds.) *Logical Aspects of Computational Linguistics*. LACL 2009, pp. 95–109. Springer, Berlin (2001)
6. Casadio, C., Lambek, J.: An Algebraic Analysis of Clitic Pronouns in Italian. In: De Groote, P., Morrill, G., Retoré, C. (eds.) *Logical Aspects of Computational Linguistics*. LACL 2009, pp. 110–124. Springer, Berlin (2001)
7. Casadio, C., Lambek, J. (eds.): *Recent Computational Algebraic Approaches to Morphology and Syntax*. Polimetrica, Milan (2008)
8. Ghanbarpur, R.: *Pregroup Analysis of an Advanced Fragment of Persian Grammar*. M.Sc. Thesis, Sharif University of Technology, Tehran, Iran, April (2009)
9. Lambek, J.: The Mathematics of Sentence Structure. *American Mathematics Monthly* 65, 154–169 (1958)
10. Lambek, J.: Type Grammar Revisited. In: Lecomte A. et al. (eds.) *Logical Aspects of Computational Linguistics*, LNAI vol. 1582, pp. 1–27. Springer (1999)
11. Lambek, J.: Type Grammars as Pregroups. *Grammars* 4(1), 21–39 (2001)
12. Lambek, J.: *From Word to Sentence. A Computational Algebraic Approach to Grammar*. Polimetrica, Monza (MI) (2008)
13. Lambek, J.: Exploring Feature Agreement in French with Parallel Pre-group Computations, *Journal of Logic, Language and Information* 19, 75–88 (2010)
14. Moortgat, M.: Categorical Type Logics. In: van Benthem, J., ter Meulen, A. (eds.) *Handbook of Logic and Language*, pp. 93–177. Elsevier, Amsterdam (1997)
15. Ryding, K. C.: *A Reference Grammar of Modern Standard Arabic*. Cambridge University Press, Cambridge (2005)
16. Sadrzadeh, M.: *Pregroup Analysis of Persian Sentences*. in [7].
17. Yetter, D. N.: Quantales and (non-Commutative) Linear Logic. *Journal of Symbolic Logic*, 55 (1990)