



The Semantics of Paranumerals

Francis Corblin

► **To cite this version:**

| Francis Corblin. The Semantics of Paranumerals. 2005. <ijn_00000650>

HAL Id: ijn_00000650

http://jeannicod.ccsd.cnrs.fr/ijn_00000650

Submitted on 23 Nov 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Semantics of Paranumerals*

To appear in :

Indefiniteness and Plurality (eds. Tasmowski & Vogeleer), Benjamins, Linguistics Today Series

Francis Corblin

Université Paris-Sorbonne & Institut Jean Nicod (CNRS)

Abstract:

The paper contrasts two semantic subclasses among expressions combining with numerals exemplified respectively by *at least* and *more than*, and contrasts these expressions with bare numerals. Even if truth conditions are often close, dynamic properties, especially anaphora and apposition, give a basis for distinguishing numerals, numerical comparatives (*more than*), and set comparators (*at least*). The paper makes the following claims : 1) bare numerals introduce in the representation a set of exactly n individuals; 2) “numerical comparatives” (*more/less than n*) only introduce in the representation the maximal set of individuals Σx satisfying the conjunction of the NP and VP constraints, and compare the cardinality of this set to n ; “set comparators” (*at least/at most*) introduce two sets in the representation, Σx , and a witness set, the existence of which is asserted, which is constrained as a set of n Xs, X being the descriptive content of the NP.

The paper is presented in the framework of Discourse Representation Theory and is based on French data.

1. Introduction

Besides *numerals* (*one, two, three...*), there is a set of expressions which combine very nicely, if not exclusively, with numerals, like: *at least, at most, exactly, more than, less than*. Let us call these expressions *paranumerals*; this term is purely descriptive and does not contain any implicit analysis of these expressions, for which we cannot rely on a received terminology. Similar expressions exist in many languages and for the sake of illustration I will use in this paper French data including : *au moins* (at least), *au plus* (at most), *exactement* (exactly), *plus de* (more than), *moins de* (less than) *en tout* (in all).

Paranumerals can combine with numerals, but, as a rule, they do not combine with quantifiers and indefinites: *plus de deux* (more than two), *moins de trois* (less than three), *exactement quatre* (exactly four)/ **plus de plusieurs* (more than several), **au moins peu* (at least few). For some indefinites, the behaviour of paranumerals is not regular: **exactement quelques* (exactly some), *plus de quelques* (more than some), *au moins quelques* (at least quelques), *au plus quelques* (at most some). Some paranumerals combine with definites, others do not¹ :

- (1) a. J’ai invité au moins Pierre et Jean.
‘I invited at least Peter and John.’
- b. J’ai invité Pierre et Jean au plus.
‘I invited at most Peter and John.’
- c. Au moins Pierre et Jean sont venus
‘At least Peter and John came.’
- d. *Plus que Pierre et Jean sont venus
‘More than Peter and John came.’
- e. *Moins de Pierre et Jean sont venus.
‘Less than Peter and John came.’

The semantic literature, in particular the algebraic approaches of Generalized Quantifier Theory (Barwise & Cooper 1981) and Boolean semantics (Keenan & Stavi 1986) generally tends to take paranumerals as having the same behavior and to see them as close to numerals; this view is based on the following list of truth-conditional equivalences:

Un (one) \Leftrightarrow *au moins un* (at least one)
plus de n (more than n) \Leftrightarrow *au moins n+1* (at least n +1)
n exactement (n exactly) \Leftrightarrow *n en tout* (n in all) \Leftrightarrow *n au plus et n au moins* (n at most and at least)

In this paper, I will try to substantiate the following claims :

1) the semantics of (bare) numerals and numeral+paranumerals should be sharply contrasted, in particular the difference between *n* and *at least n*;
 2) there are at least two different kinds of paranumerals exemplified respectively by :

- *plus de* (more than)
- *au moins* (at least)

3) it is difficult to say for certain if *exactement* (exactly) is of the *more than* kind or of the *at least* kind, but it can be established that *en tout* (in all) and *exactement* (exactly) behave differently.

The two first claims will be discussed in details in this paper, and for the third one we will only introduce the discussion.

The general perspective adopted here for analyzing the semantics of these expressions combines truth-conditional semantics (which tends to see these expressions as close) and a deeper exploration of their dynamic properties, based on two kinds of data:

A. the interpretation of definite anaphora to these expressions, an exploration initiated by Kadmon (1987) and illustrated by examples like (1) :

(1) Pierre invitera au moins deux personnes. Il les recevra dans l'entrée.
 'Peter will invite at least two persons. He will receive them in the hall.'

In such examples the problem is to determine how the pronoun is interpreted : as a reference to a set of exactly two persons, or as a reference to the maximal set of invited persons.

B. the interpretation of appositions, exemplified by examples like (2) :

(2) Pierre invitera au plus deux personnes : son père et sa mère.
 'Peter will invite at most two persons : his father and his mother.'

To some extent, apposition data can be used to elucidate the problem raised in (A) : it seems that apposition is often interpreted as an enumeration of the set introduced by the expression it is appended to.

I will argue that in order to understand correctly the semantics of the expressions of form (*paranumeral*) *nAB*, one must distinguish two sets:

- the maximal set of individuals satisfying the intersection : $E_{\max} = A \cap B$
- a set of exactly *n* members : E_n

More precisely the claim is that what distinguishes these expressions is the nature of the set(s) relevant for the computation of their representation. The relevant sets are given in the following table, in which we introduce a working terminology for the different categories we wish to distinguish:

Expression	relevant set(s)	example
Numerals :	E_n	three boys
Numerical comparatives :	E_{\max}	more than three boys
Set comparators :	E_n, E_{\max}	at least two boys

The paper gives empirical arguments based on the dynamic properties of these expressions for supporting the claim, and proposes a semantics based on the relevant sets which is formulated in the Discourse Representation Framework (Kamp & Reyle 1993). The paper is grounded on the “two set” analysis introduced in Corblin (to appear) for expressions of the *at least* paradigm (set comparators). I will not repeat here some details, discussions and references that the interested reader may find in this paper, in order to keep the focus of the present work on a contrastive approach.

2. The semantics of (bare) numerals. A quick view.

The classical truth-conditional analysis² of sentences using a bare numeral n in a structure nAB is as follows :

$$(3) \quad [[n AB]] = 1 \text{ iff } |A \cap B| \geq n$$

It holds that the nAB sentence is true if and only if the cardinal of the intersection of A and B is at least n .

It is, in general, supplemented by the classical pragmatic Gricean *implicature*, (which does not hold for *at least n* sentences), stating that the speaker has no evidence that $|A \cap B| > 1$.

$$(4) \quad n = |A \cap B| \quad \text{From (3) and the Gricean implicature.}$$

The approach of bare numerals as indefinites adopted by Discourse Representation Theory admits (3), is agnostic about (4), but holds that the statement $n AB$ "introduces" for the following discourse a set of exactly n members, $E_n \subseteq E_{\max}$.

Consider for instance (5) :

$$(5) \quad \text{Deux étudiants ont appelé.}$$

‘Two students called.’

(5) is considered true if more than two did, but cannot be followed by (6) :

$$(6) \quad \# \text{ Ces trois étudiants étaient Pierre, Jean and Nicole.}$$

‘These three students were Pierre, Jean and Nicole.’

The ingredients of the solution for accommodating these data are represented in

(7) :

(7)

X
student (X)
$ X = 2$
called (X)

Truthful embedding iff $E_{\max} = |S \cap C| \geq 2$

Accessible for anaphora $E_n : E_n \subseteq E_{\max}, |E_n| = 2$

with S for students, C for callers.

It amounts to defining the truth of (7) by means of a truthful embedding, and to take anaphora as a clue about what is made accessible by the sentence for a later reference.

The kind of dynamic data subsumed by (8) ,

- (8) Deux étudiants ont appelé... Ces deux (#trois) étudiants...
 ‘Two students called... These two (#three) students...
 can be interpreted roughly as follows.

The demonstrative NP *These n students* must be identified to a previously introduced discourse referent (DR), and its descriptive content *n students* must be satisfied by this DR. So, if the descriptive content of the demonstrative is "exactly" *n*, then its antecedent DR must be exactly *n*. That *These n Ns* means *exactly n Ns* seems to be established by the falsity or unacceptability of (9) and by the fact that (10) is a tautology :

- (9) *Ces deux étudiants sont Pierre, Jean et Nicole.
 ‘These two students are Pierre, Jean and Nicole.’

- (10) Ces *n Ns* sont $\{a, \dots\}_E \rightarrow |E| = n$

It must be concluded that *These n Ns* refers to exactly *n Ns*, which establishes that *n Ns* introduces exactly *n Ns*.³

In the DRT approach, it is thus the (bare) numeral itself which introduces a set of exactly *n Ns*. This is a very important difference with Kadmon (1987) and Evans (1980), which makes the definite NP itself responsible for the unique (or maximal) interpretation of the anaphoric definite NP.

The apposition data seems to show that it is actually the numeral which is relevant. In sentences where a list of proper names is appended to an *nAs* NP, sentences with a list of more than *n* names is not acceptable. Sentences with a list of less than *n* names is not acceptable, except with an intonative marking indicating clearly that the enumeration is not exhaustive.

- (11) Deux personnes sont venues : *Pierre, Jean and André.
 ‘Two persons came : *Pierre, Jean and André.’
 (12) Est-ce que deux personnes sont venues? Oui. *Pierre, Jean et André.
 ‘Did two persons come? Yes : *Pierre, Jean and André.’

Anaphora and apposition show, in other words, that the semantics of bare numerals in simple episodic sentences *n AB* introduces a set of exactly *n* members satisfying A and B.

3. Numerical comparatives: *plus de, moins de. [more than, less than]*

We use, as a working terminology, the label “numerical comparatives” for the French equivalent of *more than n, less than n*.

- (13) Plus de cinq personnes sont venues.
 ‘More than five persons came.’
 (14) Moins de vingt étudiants se sont inscrits.
 ‘Less than twenty students registred.’

According to the classical view, an expression like *more than n* is true is the same models than the numeral expression *n+l*.

- (15) J’ai écrit trois articles \leftrightarrow J’ai écrit plus de deux articles.
 ‘I wrote three papers \leftrightarrow ‘I wrote more than two papers.’

This sounds roughly correct, at least if one operates only with integers on the considered domain; if for instance one is allowed to consider fractions, the equivalence does not hold, as exemplified by (16) :

(16) Cela mesure trois kilomètres. \neq Cela mesure plus de deux kilomètres.

‘It is three kilometers long.’ \neq ‘It is more than two kilometers long.’

Since one can consider, say two kilometers and a half, it is not true that being more than two kilometers long implies being three kilometers long.

But the semantics of numerals and par numerals are different. It can be shown that numerical comparatives do not introduce a set of exactly n members, as numerals do (see §2 above). Consider, for instance, the contrast between (17) and (18) :

(17) Deux personnes ont été contactées : Jean et Nicole.

‘Two persons have been contacted : Jean and Nicole’.

(18) Plus de deux personnes ont été contactées : *Jean et Nicole.

‘More than two persons have been contacted : Jean and Nicole’.

(17) is fine for everyone, but (18) is awkward for most speakers.

This contrast, based on apposition, is confirmed by anaphora : (19) is correct, but (20) is not :

(19) J'ai lu trois articles. Ces trois articles sont A, B et C.

‘I read three articles . These three articles are A, B, and C.’

(20) J'ai lu plus de trois articles. *Ces trois articles sont A, B et C.

‘I read more than three articles . These three articles are A, B, and C.’

In the line of the interpretation of these facts adopted before, we can conclude that numerical comparatives do not introduce any set E_n of cardinality (*exactly*) n .

Moreover, data indicate that numerical comparatives do introduce the maximal set E_{\max} in the semantic representation. The following discourses are perceived as natural by many speakers :

(21) J'ai cité plus de deux auteurs : Platon, Aristote et Sénèque.

‘I mentioned more than two authors : Plato, Aristotle and Senecue.’

(22) J'ai cité Chomsky plus d'une fois : une dans l'introduction, une dans le chapitre 1, ...

‘I mentioned Chomsky more than once : one time in the introduction, one in chapter 1,...’

In (21) with a conclusive intonation and the presence of the "et", the list is interpreted as exhaustive. In (22) it is only required that the list be of cardinality $n+1$.

It is fair to say that some speakers do not like sentences like (21), but anaphora confirms that E_{\max} is actually part of the picture:⁴

(23) Elle a reçu plus de dix lettres, les a lues et classées.

‘She received more that ten letters, she read and filed them.’

All the speakers I asked said that in (23) she read and filed all the letters she received. The same is true for (24), an example involving “moins de” :

(24) Il a fait moins de cinq fautes . Il les a corrigées.

‘He made less that five mistakes. He corrected them.’

= He corrected all the mistakes he did.

We can conclude that for the representation of numerical comparatives, E_{\max} , and only E_{\max} is implied.

The fact that E_{\max} is not relevant for numerals (see above) but is relevant for numerical comparatives is a confirmation (contra Evans) that it is the nature of the antecedent expression itself, not the definite anaphoric NP which is responsible for the maximal interpretation. This is confirmed by the similar behavior of anaphora and apposition (in which there is no definite NP to be interpreted).

In order to represent the maximal set, I will make use of the abstraction operator of Kamp & Reyle (1993), noted “ Σx ”. The abstraction operator is associated to a subordinate DRS, and returns the set of all individuals (if there is any) satisfying this DRS. Although my representation of *more than n* is very close to Kamp & Reyle’s representation (on page 455), my approach differs on two points :

I propose this representation only for the *more than* paradigm, not for the *at least* paradigm;

I use only one RD, Σx , and not two (Σ/η).

Using this notation, the following DRS can be proposed as a correct semantic representation for *plus de deux (more than two)*. The contrast to the representation of the bare numeral is recalled in the table :

<p>Deux étudiants ont appelé ‘Two students called’</p> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 10px auto;"> <p>X student (X) called (X) $X = 2$</p> </div>	<p>Plus de deux étudiants ont appelé ‘More than two students called’</p> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 10px auto;"> <p>Σx $\Sigma x : x$ <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 5px auto;"> <p>student (x) called (x)</p> </div> $\Sigma x > 2$</p> </div>
--	--

Intuitively, the proposed representation for *more than n* :

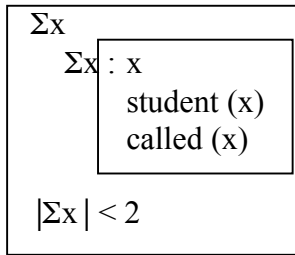
- 1) states that the set satisfying the intersection has more than two members;
- 2) *introduces* this maximal set in the representation making it available for anaphoric or appositive references.

Discourse Referents of type Σ have some peculiarities, as compared to the standard DRs, that is to say atomic and plural DRs:

1. any maximal set is unique;
2. there is no claim (in general) that such a set, built by abstraction on properties, exists.

The first property means that any truthful embedding will project a given Σ on the same set of individuals, and the second property is required by the decreasing operators like *less than n*.

- (25) Moins de deux étudiants ont appelé.
‘Less than two students called.’



In (25) the presence of a DR at the top-level does not imply that the sentence is true iff we can assign individuals of the Model to this DR. The sentence means that Σx , if not empty, has no more than two members.

From the representation assigned to such expressions, it is possible to derive some pragmatic constraints on their use. They are analyzed as comparisons : the cardinal of the maximal set is compared to a number. As for any comparison, the speaker should have a reason to evaluate a cardinal by a comparison to this particular number without giving the cardinality of the set. There are at least two good reasons one can imagine :

1. Because in the context there is a statement, or an expectation that the cardinality is n . Consider, for instance, cases in which n is a threshold. If such a standard of comparison is provided in the context, it is even possible to give the comparison *and* the cardinality of the set :

(26) Plus de cinq personnes sont venues. Elles étaient en réalité huit.

‘More than five persons came. They were actually eight.’

2. Because n is a round number.

(27) Il y a plus de cent inscrits dans ce groupe

‘There are more than one hundred registered people in this group.’

If none of these conditions holds, it is likely that the sentence will be odd, as (28) for instance is :

(28) Il y avait plus de 187 personnes à la réunion.

‘There was more than 187 persons at the meeting.’

Any speaker accepting (28) will do so, 187 being not a round number, because she thinks condition 1 holds, 187 being, for some reason, a good standard of comparison, or an expectation.

The satisfaction of one of those conditions answers the question : why do you give this comparison without *giving* the cardinal itself, if you get it?

Another question is : how can you be sure that this comparison is correct without *knowing* the cardinal of the actual set itself? Many situations realize this condition : suppose you begin to count, and then stop at n for some reason, although you are aware that there are some other cases remaining to be counted. It is then natural to state : “All I can say for sure is that there are more than n Xs.”. Another situation is the following : you know, by counting, that a given set A has the cardinality n , and by a rough comparison that another set B is smaller : it is then natural to say that B has less than n elements.

Another typical answer to the question is : it is impossible that the actual set be (bigger/smaller), because n is a threshold. Consider for instance : John has a driving license. He is more than 18, then”.

There are also obvious constraints on anaphora and apposition which come from the special nature of Σ : for instance, *less than n* can be satisfied in models in which

there are zero, one, or more individuals of the specified kind; this makes the use of a pronoun insecure, because it might be the case that there is no corresponding referent at all, and in cases there is, one cannot decide if it is atomic or plural. For apposition, it is plausible that some speakers are reluctant to interpret a list as the exhaustive enumeration of Σx because the first part of the sentence does not give the cardinality of the introduced set. In my view, the same problem arises for vague plural indefinites :

- (29) J'ai lu des livres : A, B, and C.
 'I read some books : A, B, and C.'

For such sentences, it is not clear whether the speaker gives a sample, or gives the entire list. This is probably why some speakers do not like apposition to numerical comparatives.

In other words, the specific nature of the postulated set Σ can explain why sentences with numerical comparatives have a restricted dynamic potential.

3. Set comparators : *au moins, au plus [at least, at most]*

There are some differences between set comparators and numerical comparatives :

- A. Set comparators are floating expressions:
 (30) Au moins deux personnes sont venues.
 Deux personnes au moins sont venues.
 Deux personnes sont venues, au moins.
 'At least two persons came.'
- B. Set comparators combine with cardinals, but also with definites :
 (31) Il a invité au moins ses parents et ses frères.'
 'He invited at least his parents and his brother.'
- C. As recalled by Krifka (1999) they combine with nominal predicates denoting degrees on a scale.
 (32) Si cette dame est flic, l'est au moins générale. B. Lapointe
 'If this woman is a cop, she is at least a general.'
- I will leave aside, in this paper, the uses of these expressions as discourse particles exemplified in (33) :
- (33) Mais au moins, le travail était fait.
 'But at least, the work was done.'
- This means that I will concentrate on expressions having scope over an NP.

Kadmon (1987) was the first to note that expressions like *at least* provide two interpretations for a pronoun : a reference to the maximal set and a reference to a set of (exactly) *n* elements.

- One might add that this extends to definite or demonstrative anaphoric NPs :
 (34) Au moins deux personnes sont passées ici. Ces deux personnes ont laissé leur trace.

'At least two persons came here. These two persons left their footprints.'

In (34), using the results of the argument given in §1 above, one must conclude that the first sentence introduces a set of *exactly two* elements. But the same first sentence can provide another interpretation, illustrated by (35):

(35) Au moins deux personnes sont passées ici. Ces personnes, dont nous n'arrivons pas à déterminer le nombre exact (deux seulement, trois, quatre, etc.), ont fait du feu.

'At least two persons came here. These persons (we cannot state for sure how many actually came) made a fire.'

In (35), as made explicit by the parenthetical comment, the demonstrative refers to the maximal set of individuals satisfying the conditions expressed by the first sentence.

Data involving an apposition to *at least* expressions confirm this :

(36) Au moins deux personnes sont passées ici : Jean et Pierre.

'At least two persons came here : Jean and Pierre.'

In (36), the appended expression refers to a set of exactly two persons. This expression comes usually with a falling intonation suggesting that this expression is the exhaustive enumeration of some set. In the logic of the present analysis, we are led to conclude that a discourse referent for this set is introduced in the previous sentence. But sentences like (37) are acceptable as well :

(37) Au moins deux personnes sont passées ici : Jean, Pierre, Nicole...

'At least two persons came here : Jean, Pierre, Nicole...'

The intonation is usually rising, and the sentence suggests then that the list is not finished.

(36) is perfectly natural for all speakers ; (37) is sometimes found less natural, but it is very often accepted. It is also possible to find a list of more than *n* elements, with a "et" prefixed to the last element, as in (38), the list being considered as the exhaustive set of individuals satisfying the predicates of the sentence :

(38) Au moins deux personnes sont venues : Jean, Pierre et Nicole.

'At least two persons came : Jean, Pierre and Nicole.'

From these observations, we can draw the following conclusions:

1. *at least n* can introduce in the representation a set of (exactly) *n* elements, (like numerals);
2. *at least n* can introduce in the representation the maximal set (like numerical comparatives).

The most intriguing point is the one supported by the strongest empirical data, that is to say the fact that *at least n* introduces a set of exactly *n* elements (see (34) and (36)).

There are two solutions for accommodating the accessibility of these two sets :

A . The expression *at least n* is ambiguous. Considering that a lexical ambiguity is not likely, the source might be a syntactic ambiguity of *at least*. This is the view adopted by Kadmon (1987) which postulates that *at least* can be either the modifier of the numeral determiner giving a complex determiner *at-least-n*, or an expression taking a whole NP in its scope (*at least (n Ns)*). If *at least* is analyzed as a complex determiner, the sentence introduces the maximal set; if *at least* is conceived as an operator having an NP prefixed by a numeral in its scope, this NP introduces a set of exactly *n* members, as the bare numeral does in isolation.

I discussed in details the problems raised by this approach in Corblin (to appear).

B. The relevance of two sets for the semantic representation of *at least* is not a matter of ambiguity, but a basic component of the semantics of the lexical expression : this expression introduces two sets abstracted over the syntactic environment, and states that a given relation holds between these sets. In this view, each occurrence of the expression makes these two sets accessible.

In this paper, I explore the B approach.

The central idea is that *at least/at most* introduce two sets, the maximal set Σx , and a set of cardinality n , and compare the cardinality of these sets.

The simplest way to get the semantics of *at least* from what have been assumed for n and *more than n* so far, would give the following contrast:

<div style="border: 1px solid black; padding: 10px; width: fit-content; margin: 0 auto;"> X student (X) called(X) $X = 2$ </div>	<div style="border: 1px solid black; padding: 10px; width: fit-content; margin: 0 auto;"> Σx $\Sigma x : x$ <div style="border: 1px solid black; padding: 5px; display: inline-block; margin: 5px;"> student (x) called (x) </div> $\Sigma x > 2$ </div>	<div style="border: 1px solid black; padding: 10px; width: fit-content; margin: 0 auto;"> $X, \Sigma x$ $\Sigma x : x$ <div style="border: 1px solid black; padding: 5px; display: inline-block; margin: 5px;"> student (x) called (x) </div> student (X) called (X) $X = 2$ $\Sigma x \geq X$ </div>
Deux étudiants ont appelé ‘Two students called’	Plus de deux étudiants ont appelé ‘More than two students called’	Au moins deux étudiants ont appelé ‘At least two students called’

To have two sets in the representation may seem to give a chance to accommodate the dynamic properties : for instance, in (39), the expression *Pierre et Jean* is an exhaustive enumeration of the set X (a set of exactly n members) :

- (39) Deux étudiants au moins ont appelé : Pierre et Jean.
 ‘Two students at least called : Pierre and Jean.
 $X = \{\text{Pierre, Jean}\}$

In (40), by contrast, the expression *Pierre, Jean et Marc* is an enumeration of the maximal set Σx :

- (40) Deux étudiants au moins ont appelé, Pierre, Jean et Marc.
 ‘Two students at least called : Pierre, Jean, and Marc.’
 $\Sigma x = \{\text{Pierre, Jean, Marc}\}$

But the first difficulty is to state how these sets are constrained. The semantics of the postulated representation for *at least* given above is awkward: it states that *there is* a set of n elements satisfying a given set of conditions S , *and* that the maximal set satisfying S contains n elements or more than n . Note that if one considers simple sentences expressing separately these information, it is impossible to relate them with a conjunction :

(41) Il a écrit deux livres et * Il a écrit deux livres ou plus de deux livres.

‘He wrote two books *and he wrote two books or more than two.’

The only ways to combine these information are (42) or (43) :

(42) Il a écrit deux livres ou plus de deux livres.

‘He wrote two books or more than two books.’

(43) Il a écrit deux livres et peut être même plus de deux.

‘He wrote two books and maybe even more.’

Another problem is related to the formulation of the comparison itself. In the working representation above, the comparison is between the cardinality of two sets. It is easy to show that such a definition would not apply to sentences in which the scope of *at least* is a definite NP, as in (31). What (31) means is that the maximal set of invited persons will *include* the set denoted by the NP “his parents and his brother”.

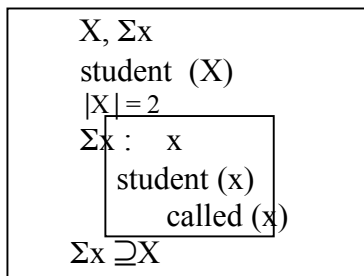
If one wants to keep the two set analysis and deal with these two problems, a solution emerges which is as follows :

1. *At least* sentences introduce two sets in the representation;
2. One of these sets is Σx , the maximal set of individuals satisfying the set of conditions expressed in the sentence (except cardinality).
3. The second set, X , is constrained by the sole NP in the scope of *at least*.
4. *At least* expresses the set theoretic relation $\Sigma x \supseteq X$.

A solution making use of these features is given in (44) :

(44) Au moins deux étudiants ont appelé.

‘At least two students called.’



This representation provides two sources for anaphora and apposition, namely Σx and X . If X receives a specific interpretation, apposition elucidates the extension of the set; X can be used without the speaker having a specific set in mind and it amounts, then, to a mere cardinality specification of Σx .

At least n and *n* are verified in the same models, although their commitments are different. The *at least* sentence deals with Σx , and specifies its extension by means of a disjunction (“ \supseteq ”), whereas in the *n* sentence, the speaker commits herself to no more than the mere existence of a set of *n* satisfiers.

Its is probably this difference which motivates the thesis that “[atleast/atmost] modifiers express modal meaning” (Geurts and Nouwen 2005), a position I adopted in the first presentation of this material. I am now less confident that this modal flavor should be considered as a part of the meaning. Although a full discussion is far beyond the scope of this paper, a few comments might be in order.

The modal component one might want to consider, states that it is *possible* that more than *n* elements verify the sentence. It seems that such a commitment could be

seen as a pragmatic inference derived from the assertion of the disjunction “ $\Sigma x \supseteq X$ ”, not as a part of the meaning proper, and this is a line of thinking I will assume in this paper.

The representation (44) mirrors the following intuition about the *at least* sentence : it introduces a set of exactly n Xs, and states that this set of Xs is a subset of the maximal set of Xs verifying the predicates of the sentence.

Dynamic data regarding anaphora and apposition show that a two sets analysis is needed for *au plus*.

The *exactly n* set is needed, as illustrated by examples like (45) :

(45) Deux personnes au plus sont venues ici : Jean et Marc.

‘At most two persons came here : Jean and Marc.’

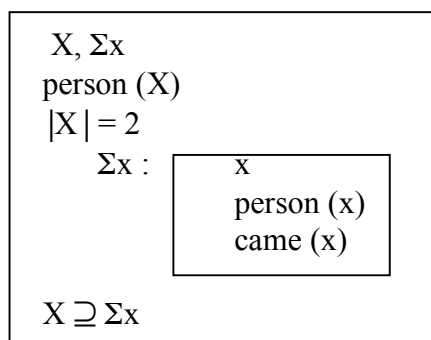
The meaning of (45) is roughly : the set of persons who came is empty or inside the set {Jean, Marc}. It is difficult, then, to accommodate (45) without assuming that the sentence introduces a set of *persons* of cardinality 2. But this set cannot be a set of persons *who came*, since the sentence asserts that the cardinality of this set is zero, one, or two.

This fact is automatically derived if one makes the rather common assumption that *au plus* belongs to the same semantic category than *au moins*, namely that it takes two arguments, one of them being the set of individuals constrained by the NP it modifies, the other one being the maximal set Σx . The difference between the two items is just that the meaning of *au plus* asserts that the relation between the two sets is : $X \supseteq \Sigma x$.

In what follows a tentative representation of *at most* sentence will be provided and exemplified as a representation of (46). In designing this representation, I wish to satisfy the following requirements: 1) preserving the minimal constraint of empirical adequacy for truth conditions and dynamic properties; 2; keeping close to the intuition which takes *at most* and *at least* as related expressions.

(46) Deux personnes au plus sont venues.

‘At most two persons came.’



As previously said, such a representation does not claim that Σx exists : discourse referents of type Σx , although located at the top-level of the DRS does not assert the existence of the set. The only strong existence claim associated with (46) is the (very weak) claim that there is a set X of two persons.

Note that there is a typical circumstance in which we use such sentences: if we know how many individuals of category X there are in a given domain, say n , it is

possible to assert correctly, for any property P , “At most n Xs P ”. If, for instance, there are three men in a room, “At most three men P ” is a tautology for any P .

The representation (46) derives correctly the data of apposition for (47): once again, it suffices to admit that the appended list is interpreted as the exhaustive enumeration of a set introduced in the first part of the sentence, namely the set X :

(47) Deux étudiants au plus ont appelé : Pierre et Jean.

‘Two students at most called : Pierre and Jean.’

$X = \{\text{Pierre, Jean}\}$

The commitments of (47) are :

(i) Pierre and Jean are students;

(ii) The set $\{\text{Pierre, Jean}\}$ is a superset of the set of students calling if there are any.

This predicts that a sentence like (48) is not interpretable :

(48) Deux étudiants au plus ont appelé : Pierre, Jean et Marc.

‘Two students at most called : Pierre, Jean and Marc.’

This representation predicts that two sets are made accessible for definite anaphora :

- X , a set of n Xs;

- Σx , the maximal, possibly null, intersection set, which in any case is of cardinality n or smaller than n .

A sentence like (48) is not interpretable, because the appended list cannot be interpreted as the enumeration of X (the cardinal of the appended set is not 2), and cannot be interpreted as the enumeration of Σx because this set is smaller than 2.

According to the view advocated in this paper, the specificity of *at least/at most* quantification can be approached as follows : these expressions work on the basis of a "witness set" X of cardinality n , the existence of which is asserted and which is provided by the NP in the scope of the expression; they assert a set theoretic relation between this set and the maximal set of individuals verifying the conditions expressed in the sentence.

This makes them different from numerical comparatives, which compare the cardinality of the (maximal) intersection set to a number (see §2)

The witness set can be specific, like the corresponding set introduced by a numeral can be. In this case, the speaker has a specific set of n individuals in mind as a witness set, which can be, for instance, enumerated by an apposition. It can also be non-specific, like for the corresponding numeral, and then, the identification by apposition is impossible.

Again it is possible to derive modal inferences from a representation so constrained. In stating that the members of Σx , if there are any, belong to a given witness set, the speaker looks like committing herself to :

- It is possible that any member of the witness set belong to Σx ;

- It is impossible that other individuals belong to Σx .

A nice feature of this solution, is that it provides a perfect analogy between the semantics of *at least* and *at most*, which is not the case for my previous proposal in Corblin (to appear)⁵.

Note that it can also explain nicely that *at most* and *at least* can be conjoined and what happens when they are conjoined as in (49) :

(49) Three students called, at least and at most

The representation will be, in short : $X, \Sigma x : |X| = 3 \wedge \Sigma x \supseteq X \wedge X \supseteq \Sigma x$. The sentence states, in a rather complicated way, that there is a witness set containing three students, and that the maximal set Σx is this set.

There are even examples discussed in Corblin (to appear) which show that the witness set can contain entities which do not satisfy the descriptive content of the NP modified by *at most*. Consider for instance, the following example :

(50) Il y a au plus deux solutions.

‘There are at most two solutions.’

The interpretation of (50) cannot be : there is a set of two solutions, and the maximal set of solutions contains two elements or less than two. What does the sentence mean? Roughly the following : there is a set of two “things” such that the maximal number of solutions, if there are any is a subset of this set. The necessity for this set of “things” is illustrated by sentences like :

(51) Il y a au plus deux solutions : combattre, ou partir.

‘There are at most two solutions : fighting, or leaving.’

The utterer of (51) is committed to the statement that fighting or leaving are “possible” solutions, and that if there are solutions, they are in this set.

This is a special example involving existence as main predication, and involving an NP denoting entities (*solutions*) which may not exist. It should be discussed more at length when considering the construction algorithm required for providing the representation we need. In the standard cases, the general form of the algorithm will be roughly as follows :

1. build a set X of n satisfiers of the NP to which *at most/at least* is attached;
2. build a set Σx as the maximal set of Xs satisfying the main predicate;
3. add the condition stating respectively $\Sigma x \supseteq X$ or $X \supseteq \Sigma x$;

Examples like (51) would require a slight modification of this algorithm.

4. Summary : a brief comparison of numerals and par numerals

I will now try to sum up and compare the main feature of the three categories .

(a) Trois étudiants ‘Three students’ <i>Numerals</i>	(b) Plus de deux étudiant ‘More than two students’ <i>Numerical comparatives</i>	(c) Trois étudiants au moins ‘At least three students’ <i>Set comparators</i>
--	--	---

(a), (b) and (c) are true in the same Models. This is general for count nouns, if it is clear that their domain can only be divided by integers.

Only (b) and (c) introduce the maximal intersection set.

(b) introduces only the maximal intersection set.

(c) introduces the maximal set and a witness set of *n elements*.

(b) and (c) are "genuine" quantifiers in the sense that the maximal set is part of the picture.

(a) do not introduce the maximal set, but only a set of n members of the intersection set, the relation of this set to the maximal set being left semantically unspecified.

A brief look at the interaction of the three forms with negation shows many differences :

The most striking data regarding negation is that set comparators (*au moins, au plus*) are incompatible with wide scope negation.

For numerals, it is easy to state the contrast between wide scope negation and narrow scope negation :

narrow scope: the sentence claims that there is a set of the relevant cardinality which does not satisfy the predicate;

Wide scope : the sentence claims that there is no set of the relevant cardinality satisfying the predicate.

Compare the following sentences in which we try to make the wide scope interpretation (*there are five glasses such that...*) unlikely :

(52) Je n'ai pas bu cinq verres de vin.

'I did not drink five glasses of wine.'

(53) Je n'ai pas bu plus de quatre verres de vin.

'I did not drink more than four glasses of wine.'

(54) Je n'ai pas bu au moins cinq verres de vin.

'I did not drink at least five glasses of wine.'

All the speakers I asked said that (54) is awkward under the wide scope negation interpretation, and cannot be accepted unless in an echo context. In contrast, (52) and (53) can be used out of the blue with the numeral out-scoped by negation.

A possible explanation in the light of the proposal would be that the existence of a witness set is necessary for computing the interpretation of set comparators. If one tries to interpret the set X in the scope of the negation, there would be no existence claim of a witness set and the interpretation would be impossible to construct.

Another very interesting piece of data concerns the selectional restrictions of predicates taking the considered expressions as argument. M. Hackl (2001) in his dissertation on comparative quantifiers (our "numerical comparatives") points to a puzzle related to the present study.

Hackl observes the following contrast between two sentences which should be equivalent in virtue of the equivalence *more than n/ at least n+1* for entities counted with integers :

(55) ?John separated more than one animal.

(56) John separated at least two animals.

Roughly speaking, the problem is that although (55) takes *one* as a part of a complex expression "corresponding" to a plurality (two, three, or more), it appears as if the only relevant feature for selection were the offending singular "one", exactly as in **John separated one animal*.

In the line of our hypothesis what is the problem like?

At first glance, our theory predicts that for *at least n*, *n* should satisfy the selectional requirement exactly as a bare numeral should do. The reason is that we have, so to speak, a numeral interpretation "within" the representation of the complex expression. Prediction (57) is borne out, and illustrated by (58) :

(57) If *n* violates a selectional restriction, *at least n* violates it.

(58) John separated *one animal → John separated *at least one animal.

But it is fair to say that our approach does not expect any problem with the *more than* case. The representation contains only the maximal interpretation Σx , which stands for a set of a cardinality which should satisfy, in cases like (55), the selectional requirement.

Our own theory predicts that there should be a strong contrast between:

John separated at least one animal prediction : out

John separated more than one animal prediction : correct

It seems to me that there is actually a strong contrast, and that the *more than* cases are not as bad as Hackl suggests; it can even be observed that there are in French some colloquial uses of "plus d'un" (*more than one*) as an understatement for *many*. I

think that such expressions can be used even if the predicate imposes a plurality constraint on its argument.

(59) Ce chronomètre très précis a départagé plus d'un concurrent.
'This very precise chronometer decided between more than one concurrent.'

(60) Ce surveillant est amené à séparer plus d'un élève dans une journée.
'This supervisor has to separate more than one pupil within a day.'

To my judgment, these examples are correct, although their "at least" correspondent would be bad.

5. Some paranumerals difficult to classify

Other expressions like *exactement*, *à peu près*, *environ*, *en tout*, fall under the descriptive concept "paranumeral" used as a working notion at the beginning of this paper.

I will consider each of these expressions trying to see if it falls easily in one of the two categories considered up to now.

Exactement [Exactly].

It is difficult to state if *exactement* should sit on the *more than n* side, or on the *at least n* side. The problem is that its lexical meaning makes the two postulated sets, distinguished in the course of this study, identical. It is thus difficult to use dynamic data for establishing which set is made available.

The behavior with negation looks more similar to what happens for *more than n* : narrow scope interpretation under negation is acceptable as illustrated by (61) :

(61) Je n'ai pas parcouru exactement deux kilomètres.
'I did not walk two kilometers exactly.'

But it is a weak argument for deciding, and I leave the question open.

Environ, à peu près [about]

It is difficult in this case to use dynamic data for contrasting two sets because of a complication : *environ* and *à peu près* select round numbers. For instance, (62) is strange because 47 is not a round number :

(62) *J'ai à peu près 47 étudiants dans mon cours.
'I have about 47 students in my course.'

For this reason, it is difficult to contrast an "exactly *n*" interpretation to a maximal one.

Again, I think that the present discussion does not give any conclusive argument for choosing to see *environ* and *à peu près* as closer to *more than n* than to *at least n*.

En tout [in all]

Although *n exactement* and *n en tout* are very often equivalent, there are some good reasons for doubting that they belong to the same category.

en tout n can only be used if *n* is obtained by a cumulation of different numbers :

(63) Pierre mesure 1,80 mètres exactement (*en ⁶tout).
'Pierre is 1,80 meters tall exactly (in all).'

En tout n preserves the possibility to interpret *n* as a round number, which *exactement* prohibits.

- (64) J'ai cinquante étudiants exactement. (False if I have 49 students)
 'I have fifty students exactly.'
- (65) J'ai cinquante étudiants en tout. (True if I have 49 students under the interpretation "round number" of fifty).
 'I have fifty students in all.'

So *en tout* does not belong to the *exactement* paradigm. It is probably another kind of paranumeral which is associated to the notion of cumulation.

6. Conclusion.

The main aim of the paper was to contrast different semantic subclasses among paranumerals and to show that many expressions which are often put together should be carefully distinguished.

Even if truth conditions are often similar, dynamic properties, as shown in the paper, give a basis for distinguishing numerals, numerical comparatives, and set comparators. Numerals introduce in the representation a set of exactly n individuals, satisfying the conjunction of the NP and VP constraints and assert the existence of this set. Numerical comparatives (*more/less than n*) only introduce in the representation the maximal set of individuals Σx satisfying the conjunction of the NP and VP constraints, and compare the cardinality of this set to n . Set comparators (*at least/at most*) introduce two sets in the representation, Σx , and a witness set, the existence of which is asserted, constrained as a set of n Xs, X being the descriptive content of the NP.

The paper focuses on typical properties and suggests the main features of DRT representations dealing with these properties for a restricted set of distributions, namely the use of those expressions in construction with NP containing a numeral.

Many details have been left aside for trying to give a survey of the main contrastive features of numeral/paranumeral landscape.

Among the important points to be considered carefully is the range of constructions allowed for each type of expression. It remains to be explained, after all, why those expressions can be seen as paranumerals (the association with numerals is typical), and why they are not restricted to numerals. The present approach, in other words, would have to be subsumed under a more general theory, powerful enough for applying to the all range of distributions, and for explaining why numerals are a typical distribution. Such a theory is far beyond the scope of this paper, but I would like to suggest, as an opening, a contrast in line with the difference postulated in this paper.

At least/at most cannot combine with degree adjectives:

- (66) Il est au moins *grand /*froid
 'It is at least great/cold'

At least/at most can only work if they are attached to a constituent, which can be interpreted as identifying a precise measure on a scale. The constraint is possibly less absolute for comparatives, although it applies strictly to decreasing comparatives :⁷

- (67) Il est moins que *grand /*froid.
 'It is less than great/cold.'

A detailed study of the restrictions would be in order before deciding if this requirement holds only for *at least/at most*, or can be generalized.

The notion of "witness set" used in the present approach for explaining the kind of semantic calculus associated to *at least/at most* would provide a good explanation of why this is so : a witness set works as some sort of yardstick, used for the evaluation

of a measurable dimension of the maximal set; this is why an acceptable argument of *at most/at least* must provide a definite point on a scale (a cardinality, or the name of a recognized degree) and not a vague comparison to a standard, as degree adjectives would do. More work on the *more than* paradigm would introduce, if the requirement can be generalized, a very interesting contrast between *more +than+adj*, which put a ban on degree adjectives, and *more+adj+than...* which selects a degree adjective. The fundamental contrast to be explored is the contrast between comparing the dimension of something to a measure on the relevant scale, and comparing the location of two entities on a scale.

References

- Barwise, J. and Cooper, R. 1981. "Generalized quantifiers and natural language". *Linguistics and Philosophy* 4:159-219.
- Corblin, F. 1997. "Les indéfinis: variables et quantificateurs". *Langue Française* 116: 8-32.
- Corblin, F. 2002. *Représentation du discours et sémantique formelle : Introduction et applications au français*. Paris, P.U.F.
- Corblin, F., and de Swart, H. eds .2004. *Handbook of French Semantics*, Stanford, CSLI Publications,
- Corblin, F. to appear. "Existence, maximality, and the semantics of numeral modifiers", *Existence*, Comorovski I. and von Heusinger K. eds. Kluwer. *In prep.*
- Dupuy, P. 2004. "En tout et exactement", MA. thesis, Paris-Sorbonne.
- Evans, G. 1980. "Pronouns", *Linguistic Inquiry* 11, 2: 337-362.
- Hackl, M. 2001. *Comparative quantifiers*, Phd. MIT.
- Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Amherst: University of Massachusetts.
- Heim, I. 1990. "E-type pronouns and donkey anaphora". *Linguistics and Philosophy* 13: 137-177.
- Geurts, B. and Nouwen, R. 2005. *At least et al.*, ms.
- Kadmon, N. 1987. *On Unique and Non-Unique Reference and Asymmetric Quantification*, Amherst: University of Massachusetts.
- Kamp, H. and Reyle U. 1993. *From Discourse to Logic*. Dordrecht, Kluwer.
- Keenan, E. and Stavi J. 1986. "A Semantic Characterization of Natural Language Determiners", *Linguistics and Philosophy* : 254-326.
- Krifka, M. 1992. "A compositional semantics for multiple focus construction", In J. Jacobs (ed.) *Informationstruktur und Grammatik*. Opladen: Werdeutscher Verlag, 17-53.
- Krifka, M. 1999. "At least some determiners aren't determiners", in K. Turner (ed.) *The semantics/pragmatic interface from different point of views*. Elsevier, 257-291.
- Landman, F. 2000. *Events and Plurality: The Jerusalem Lectures*, Dordrecht: Kluwer.
- Laserson, P. 1999. "Pragmatic Halos". *Language* 75.3 : 522-552.
- Merin, A. 2003. "Replacing "Horn Scales" by Act-Based Relevance Orderings to Keep Negation and Numerals Meaningful". <http://semanticsarchive.net/>
- Partee, B. 1987. "Noun Phrase Interpretation and Type Shifting Principles", Groenendijk et al. (eds) *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Foris : Dordrecht, 115-143.
- Rooth, M. 1985. *Association with focus*, PhD. dissertation, Amherst.
- Sells, P. 1985. *Restrictive and Non Restrictive Modification*. Report 85-28. CSLI : Stanford.
- Stanojevic, V. 2004. *Syntaxe et sémantique des noms de nombre*, Thesis, Université Paris7- Université de Belgrade.

* This paper is a version of the talk presented at the conference « Indefinites and Weak Quantifiers » held in January 2005. I owe many thanks to two anonymous readers of this paper. Their comments and demands of justification of the version presented at the conference have played a great role in convincing me that it is possible to provide a semantic representation of *at least/at most* without any appeal to a modal part of the representation (for an opposite view, see Geurts and Nouwen 2005). This explains that the paper presents an analysis of these expressions which is closer to my former analysis (Corblin, to appear) than the version presented at the conference. I read Geurts and Nouwen (2005) when I was finishing the revision of the present paper, which explains that it was not possible to incorporate a

discussion of their work. Since my first presentation of this material in the 2002 Nancy's workshop "Existence : semantics and syntax", I got very interesting comments from B. Geurts, and I had the opportunity to hear a couple of talks by him on this topic. I am now convinced that a semantics without built in modality is a better way to deal with paranumerals, and I encourage the reader to read Geurts and Nouwen (2005) for an opposite view. My work on this topic has greatly benefited from discussions with many other people, among others, G. Chierchia, P. Dupuy, O. Matushansky, A. Merin, and V. Stanojevic.

¹ I owe to an anonymous reader of this paper the following example :

Il faudrait plus que Pierre et Jean (la dernière réunion/mon dernier échec) pour me décourager.

'More than Peter and John (the last meeting, my last failure) would be required for depressing me'.

There are many differences in French between the acceptability of *plus que/moins que* and *plus de/moins de* that I will not discuss in this paper.

² A. Merin (2003) provides a very strong criticism of this dominant view, and gives very good argument for coming back to the thesis that *n* means "n", and has no other meaning.

³ For the sake of simplification, I do not take into account in this paper the contrast between round numbers and others.

⁴ A fact discovered by Kadmon (1987).

⁵ The main difference is that in my previous treatment, the set *X* in the representation of *at least* was defined as a set of individuals verifying the descriptive content of the NP *and* the predicate (very similar, then, to the interpretation of a bare numeral). Such a choice has two negative consequences: the representation of *at least* is, so to speak redundant, and the plain extension of this representation, *mutatis mutandis*, to *at most* is impossible. In this slightly different version, these two problems are fixed.

⁶ A point made by Pascal Dupuy (2004).

⁷ Positive comparatives can be used more freely, although they produce with many degree adjectives a meaning close to « very » : « C'est plus que froid » ('It is more than cold') is interpreted as : « It is very cold ».