ON THE MULTIPLE EXPRESSION OF NEGATION IN ROMANCE

Francis Corblin & Lucia M. Tovena

U. Paris IV, UMR 8528 CNRS-U. Lille 3

1 Introduction

The aim of this paper is to deal with the expression of negation in some Romance languages, and to do so by deriving the so-called “negative concord” phenomena from more general principles and constraints on the occurrence of negative expressions within the clause boundaries. Our proposal is built on two general principles:

1- A principle regarding the lexical marking of argument variables in the scope of a negation. This principle says that in a given natural language, there is a way of marking unambiguously an argument variable as having to be interpreted in the scope of clause-mate negation.

2- A principle limiting the recursivity of negation within the verb-argument domain. The consequence of this constraint is that there is, in the default case, one negation per clause. Our assumption is that these two principles are generalisations valid for most, if not all, natural languages. In order to cope with most varieties of Romance, we assume that a third principle is necessary:

3- A constraint on the distribution of negative expressions in the clause, namely “Neg-first”. This principle requires the presence of a preverbal negative expression in the surface structure of the clause.

Principle (3) is neither universal nor specific to Romance languages. It appears nevertheless to constrain a very large set of natural languages, and we will show in this paper how it should be tuned for some Romance languages (Romanian, French and Italian).

We are very grateful to D. Godard, I. Comorovski, T. Espinal, B. Laca and all colleagues who helped us with the data. Special thanks are due to H. de Swart and J. Doetjes for their comments on the Going Romance 99 presentation of the first version of this paper, and to J. Rooryck, who suggested many improvements of the final version. This project has been supported by the cooperation project CNRS-NWO "Sémantique formelle et données du français".
For each of these principles we rely on previous work, although our formulations are sometimes slightly different from the original theses.

Our main sources for principle (1) are Ladusaw (1992) and Dowty (1993). We derive principle (2) from our understanding of Ladusaw (1991) and previous work by Corblin (1996) and Tovena (1996). Our main inspirations for principle (3) are Jespersen (1917), Dahl (1979) and Horn (1989).

2 The multiple expression problem

In this paper, negation refers to a semantic notion required in the description of natural languages for which the negation of classical logic can be taken as a workable translation. A negative sentence is a sentence for which the best translation in logic will put the translation of the main verb in the scope of the logical operator for negation. A negative expression, or negative for short, is the specific lexical item responsible for this translation. In many natural languages, including Romance languages, the multiple expression problem is due to the fact that the occurrence of one negation in the translation is triggered by more than one negative. This can be seen as a problem because if the meaning of a negative is negation, then two negatives should end up as two negations. Note that combining more than one negation is not a problem for semantics, negation being a recursive operator in the semantic language used for translation. Rather, this multiple expression might be seen as a problem for compositionality. This is a problem too, if one looks at the data in terms of economy or optimality. Why should a meaning be expressed by the repetition of lexical items where one of them would be enough to do the job?


Not all these studies consider that the expression of negation occurs more than once, or that there is a compositionality problem with negation. Some proposals have tried to show that the device handling negation is not specific to it:

1) parallelism neg/WH, Haegeman (1995);
2) marking of downward positions (polarity item), Dowty (1993);
3) combining quantifiers: factorisation, incorporation, (reiteration/resumption), May (1989),
   Haegeman & Zanuttini (1991), de Swart (1999);
4) postulating special kinds of universal quantifiers, Giannakidou (1999).

This research might be on the right track. However, the problem with this kind of approach is,
in our view, that, at least up till now, it has some difficulty in handling all the data: it covers
only some cases, or it is very powerful and is far from making predictions very close to the
data.

Another research direction is the following: if one observes specific phenomena for the
expression of negation in a single clause, this is in virtue of what is expressed, i.e. negation in
a single clause. This is, in our view, the position adopted for instance in Ladusaw (1991,
this paper, we will introduce a proposal for deriving the most specific properties of the
expression of negation in single clauses which is based on the very terms of the problem:
- the expression of negation (a truth-functional operator associated with complementation);
- in the clause, i.e. in a semantic space organised by the verb-argument relation.

Let us briefly recast the problem. Consider (1) and (2) with the so-called “negative
concord” reading of (1):

(1) Mary didn’t talk to nobody
(2) Nobody came

What is the correct analysis of nobody?

1) Nobody is ambiguous. Sometimes it is negative, and sometimes not. Depending on the
analysis, this leads to a characterisation of this ambiguous lexical item as either \( \neg \exists x / \exists x \) or
\( \forall x \neg / \forall x \).

The main problem is that in itself, this assumption is not very natural. The idea that an
item is ambiguous between \( X \) and \( \neg X \) could be tolerated as an accident. But it is quite
unnatural to assume that, as a rather general rule, items which have negative meaning (and
negative morphology in this case) have also a reading in which this meaning is absent.

2) Nobody is not ambiguous, with two subcases:
- It has a negative meaning, but in the context of another item expressing negation, there is a
  “factorisation” of the negative meaning, and nobody does not express a negation of its own in
  such contexts: the processing of \( \neg \exists x \neg \exists y \) translates into \( \neg \exists x \exists y \). This view is the basis of
the factorisation analyses in the spirit of May (1989) and Haegeman & Zanuttini (1991). The problem is then: Why should we have this factorisation as a rather general rule in the case of a negative meaning, and within the clause boundary?

- It has no negative meaning. But how then should we explain that it is able to express negation in its own right when used in isolation, as in (2)?

The conclusion one could draw is that most analyses of the basic meaning of items like nobody in negative concord dialects run into problems, and must be supplemented by additional assumptions in order to reach some kind of empirical and explicative adequacy. The most important problem might not be in the choice of one of those basic analyses, but in the additional assumptions for the distribution of negatives within the clausal boundaries.

3 Negation within the clause

3.1 Ladusaw (1991, 1992)

In our view, there are (at least) three important (and distinct) parts in Ladusaw’s (1991, 1992) proposal about negative concord:

(3) Lexical expressors of negation are indefinites.
(4) Negation is a clausal feature, “‘semantically potent’ at the clause level” (Ladusaw 1991: 85) although it can be expressed on more than one position in a clause.
(5) There are configurational constraints on the distribution of negative features in the clause.

We take (4) to be the key part of the proposal, although, in our view, it has not received enough attention. What (4) says, roughly speaking, is that there is one negation per clause. This interpretation of Ladusaw's view is coherent with the “foot-feature” model of GPSG used by Ladusaw (1991) and with the way Ladusaw uses this constraint in his work: “All negative constituents spawn projection paths for [Neg], which flow together in good feature fashion so that, at the clause level there is only one, albeit multiply licensed, contribution of semantic negation” (Ladusaw 1991: 86).

The main interest of this proposal for us is that it points to a constraint restricting the expression of negation in the clause. Roughly speaking, the expression of more that one
negation within the clause boundaries should be exceptional, and in the default case, the derivation of a clause will end up with at most one negation in the semantic representation. As such, this view is not without problems. There are, at first, empirical problems, since some clauses are interpreted as double negation, and others are ambiguous between double negation and negative concord. There are explicative problems as well, since this is a stipulation which concerns only negation, and within clausal boundaries. It is difficult to think of another semantic feature having the property of being semantically potent only at the clause level, potentially having a recursive interpretation. One might think that Question, for instance, is semantically potent at the clause level, although it can be `spread’ on more than one constituent. But there is nothing related to questions, in our view, that might be compared to the recursive interpretation of negation within the clausal domain. A sentence like Who said What cannot be analysed as a question embedded in a question, as a recursive structure. Although it is (at least) rather difficult to imagine what recursive interrogation is, there is no such problem imagining what recursive negation is.

But Ladusaw's formulation is also a promising way of finding a meaningful connection between the specific properties of negative expressions, all of them being clause-bound: this is true for the postulated limitation (one negation per clause) and for configurational constraints (which apply within the clause). Pursuing the study in this direction, we will now introduce a proposal for deriving negative concord phenomena from general principles that constrain the expression of negation within the clause boundaries.

3.2 An overview of our proposal

We will try to derive the distribution and interpretation of negative morphemes within clausal boundaries by the interaction of three principles.

PRINCIPLE 1:
The lexical identification of argument-variables to be interpreted in the scope of a clause mate negation is relevant and most natural languages have lexical items that unambiguously mark a variable as negated.

PRINCIPLE 2:
In the default case, negation is not recursive in the verb-argument domain. The default case exhibits one negation per clause.
Under exceptional circumstances it is possible to embed one negation under another negation, but never more than one.


For a clause to end up negated (its verb being in the scope of a negation), in many natural languages, the presence of a pre-verbal negative is required. (general formulation).

“Neg-first” can be realised in different ways depending on particular languages.

From the conjunction of principles 1 and 2, we derive the fact that incorporation (morphological or semantic) of the negative (the expressor of negation) in the lexical terms used for negative variable marking is a natural solution. Our assumption is that principles 1 and 2 are general, and principle 3 applies to many languages (including most varieties of Romance languages, with the exception of spoken French and Catalan). The common feature of these principles is that they all mention negation and the verb-argument structure of the clause.

4. Expressing negation within the clause: Principle 1

In most semantic theories, the semantics of verbs distinguishes: an eventuality, associated with the verb itself; roles in this eventuality held by individuals, and individual time/space locations; meaning postulates relating the existence of an eventuality of type \( V \), and the existence of individual satisfiers of the roles. If there is a meeting, for instance, then there is at least an agent, a theme, a time location, and a space location of the meeting. The set of roles concerned by a meaning postulate of this kind might be called the set of obligatory arguments of the verb. These obligatory roles can be identified by the following test:

(6) \textit{Test on obligatory arguments:}

*J’ai mangé, mais rien

I ate, but nothing

If \( X \) is an obligatory argument, to make it the focus of a constituent negation results in an ill-formed sentence.
The classical semantic interpretation of negation is based on set subtraction: one has to take away from the class of possible Models those containing eventualities of a certain kind; those eventualities are characterised by their sort (associated with the verb itself) and by the characterisation of their arguments. For instance, in a sentence like *John did not meet Mary*, one has to subtract from the set of possible Models the set of Models including any eventuality of type MEETING having *John* and *Mary* as participants.

**Negative marker:** minimal or basic expression of negation

Most natural languages have a *negative marker*, a term expressing negation of any eventuality involving argument constants, or no argument at all, and thus compatible with any kind of verb.

**Sorted argument variables in the scope of negation:**

Consider for example the subtraction of the following set of eventualities:

MEETING, Mary, x: man (x), Berlin, 8 April 1999 (one of the arguments is a sorted variable); what has to be subtracted is any meeting having Mary as an agent, and an individual of the sort man as a theme, etc.

What is to be expressed?

1) That the verb is in the scope of a negation.
   i.e. take away eventualities of type V (meeting).

2) Which (if any) argument variable is in the scope of the negation.
   i.e. take away any eventuality of type V having any individual of a given sort as holder of a given role.

There are at least three options as to the lexical marking of scope dependencies in natural languages:

1. no lexical marking: a dependent variable is expressed by an expression that can also out-scope the operator;
2. dependency marking: the variable can only be interpreted in the scope of another operator of the context, but many different operators can play this role;
3. marking a dependency to a specific operator: the variable can only be interpreted in the scope of a given operator.
These three options could be illustrated respectively by: *Mary did not meet a man, Mary did not meet any man, Mary did not meet no man.*

4.1 Covert argument variables in negated clauses

In natural languages, an obligatory argument does not have to be expressed overtly. This is true for negative and positive sentences alike.

(7) Pierre mange
    Peter eats
(8) Pierre ne mange pas
    Peter does not eat

In (7) and (8), only one argument is expressed, the agent, although the verb *manger* has four obligatory argument: agent, theme, time/place location. How are unexpressed obligatory arguments interpreted?

In **positive sentences:**
1) Obligatory arguments can be interpreted by *existential closure* (as a consequence of meaning postulates; see the definition of *obligatory arguments* above).
   If (7) is true, there is a thing Pierre ate, and a time/place location for the eating event. Existential can be seen as the default option for unexpressed (obligatory) arguments in simple clauses without any logical operator.
2) Obligatory arguments can be interpreted by *anaphora* (with restrictions depending on the status of zero anaphora in a given language, and depending on the kind of argument). Location arguments, for instance, are more easily interpreted as anaphoric that theme arguments.

(9) Qu'a fait Pierre du sandwich? (?) Il a mangé
    What did Pierre do with the sandwich? He ate.
(10) Qu'a fait Pierre à huit heures? Il a mangé
    What did Peter do at eight? He ate.

In **negative sentences:**
1) Covert obligatory arguments can be interpreted as variables in the scope of negation:
A plausible interpretation of (11) is: take away any eventuality of EATING with Pierre as agent, and x, whatever x is, as theme. One could be tempted to pursue the enumeration and say: take away any eventuality of EATING with Pierre as agent, and x, whatever x is, as theme, y, whatever y is, as time location, and z, whatever z is, as space location. But things are not that simple. The preferred interpretation of (11) is, in our view, something like: at the specific time $t$, and in the specific location $l$, provided by the ongoing discourse, there was no EATING eventuality having Pierre as agent and x, whatever x is, as theme. In other words, the preferred interpretation of (11) takes the covert theme argument as a variable in the scope of the negation, and the covert time/space location as variables interpreted by anaphora. The consequence to be drawn is that the interpretation of covert arguments in negative sentences has access to at least two options (dependent/anaphoric), and that a very simple sentence like (11) can combine both. Therefore, we need to add explicitly another option:

2) Covert obligatory arguments can also be interpreted by anaphora.

We may illustrate this option by giving a context to (11):

(12) De midi à deux heures, Pierre était chez lui.
    Pierre n'a pas mangé.
    From midday to two o'clock, Peter was at home. He did not eat.

(12) illustrates a case where we find anaphoric interpretations of the time/space location variables, and a dependent interpretation of the theme variable. The most plausible interpretation of (12) is: at $t$ (from midday to two o'clock, a value provided by the context) and at $l$ (at home, again a value provided by the context) there was no eating eventuality involving Pierre as agent and x, whatever x is, as theme (dependent variable outscoped by negation). It will be enough for the present purpose to call "anaphoric" those interpretations of unexpressed argument variables which are set to a contextual value provided by the discourse context.

The problem of deciding between a negative and an anaphoric interpretation of covert obligatory arguments seems interesting in itself, but its discussion is far beyond the scope of this paper. The only relevant point for the present discussion is the following observation:
covert obligatory arguments in negated clauses can be interpreted as negated variables or as anaphoric expressions; many clauses combine both, and there is no automatic or mechanical rule stating what the intended interpretation of a covert argument is.

The consequence of the non-expression of an obligatory argument is thus, at least in principle, an ambiguity between an anaphoric interpretation and an interpretation bound by negation. So, non-expression of obligatory arguments is not an unambiguous way of expressing variables interpreted in the scope of negation (or any other logical operator).

There are other obvious limitations to the strategy leaving unexpressed the variable for the obligatory argument which is to be interpreted in the scope of a clausemate negation:

1. This strategy cannot express sorted variables.

2. It is ruled out by syntactic or lexical constraints which impose overt arguments: the subject is often obligatory, and with some verbs, so is the direct object. For instance, in a case where the presence of a lexical object is highly preferred, we can observe that the sentence stripped of any lexical object will be associated with a representation where the corresponding argument is interpreted as anaphoric:

   (13) Pierre n'a pas dit
       Peter did not say

   (13) can only have an anaphoric interpretation for the covert theme argument: *Peter did not say that*, i.e. whatever was being talked about.

4.2 The marking of negated variables

A language like French exhibits some interesting constraints on lexical marking of argument variables. Lexical items like *quelqu'un* and *quelque chose* can be used for variables interpreted by existential closure (in (14) this is the only available option) and for variables interpreted in the scope of a logical operator, an option available (and preferred) in (15) and (16):

(14) J’ai vu quelqu’un  existential closure
    I saw someone

(15) Si vous avez vu quelqu’un dites-le I saw someone  bound variable/exist. clos.
    I saw someone say so
Avez-vous vu quelqu'un?  bound variable/exist clos.
Did you see someone/anyone

But this is not true for negation. If quelqu'un and quelque chose occur in the syntactic scope of a negation, which is syntactically correct, they can only be existentially interpreted:

Je n'ai pas vu quelqu'un  Existential closure only.
I did not see someone

In other words, these lexical items cannot mark a variable interpreted in the scope of a clause mate negation, although they can do this job perfectly well for any other logical operator. Various hypotheses can be suggested to accommodate this restriction. Our intuition is that items able to mark an existentially closed variable might be incompatible with the marking of a negated variable, except when those variable markers can be interpreted as "minimizers". Hence the contrast between (17) and (18):

Je n'ai pas vu une personne  bound variable/exist. clos.
I did not see a person

In (18), une personne can be interpreted in the scope of the negation.

The preceding observations can be summed up as follows:
1) The lexical realisation of negated obligatory argument variables is sometimes necessary: there are constraints prohibiting covert arguments;
2) The lexical realisation is always relevant, since covert arguments in the scope of negation can be interpreted either as negated variables or as anaphoric expressions;
3) The realisation of a bound variable by an ambiguous indefinite (interpretable as existential or as a bound variable) is forbidden in some languages with respect to negation (as opposed to other logical operators).

From these observations, one can infer that natural languages will have lexical items unambiguously mark the fact that a variable must be interpreted in the scope of a clause-mate negation. "Unambiguously" means that the lexical items under consideration cannot be interpreted by existential closure.

4.3 A typology of items expressing negation in the clause
We can sum up as follows the typology of lexical items used for expressing negation, and the strategies available for variable marking.

Negative marker: meaning: the verb is in the scope of negation. The negative marker can be found with verbs without arguments, or with saturated or unexpressed argument variables.

Variable marker: meaning: a given argument is in the scope of a clausemate negation.

Negative variable marking:
- Negative marker
  + Non-veridical variable marker: polarity item

- Negative marker
  + Specific negative variable place holder

Specific negative variable place holders:
- Negative indefinites: can (only) mark a negative variable, but cannot express negation on their own.
- Morphological incorporation (of the negative marker to the variable marker): nobody, nessuno (N-words).
- Semantic incorporation (of the lexical meaning to the variable placeholder): personne, rien in Modern French.

It seems that for natural languages, incorporation is a highly favoured option for lexical variable marking. The notion of "N-word" itself, emerging from the observation of many Indo-European languages, reveals that morphological incorporation is a very widespread solution.

But why is it so? Incorporation presents an obvious advantage in terms of economy, because it spares one from of having a specific item for marking negated variables. Incorporation is also an option available in any syntactic position, as opposed to the polarity item technique, which seems to require structural conditions for licensing. Standard English may provide an illustration. In English, the use of a polarity item is the default option; this option is available for the object position:
I did not read any paper

But this option is not available for the subject position, which does not meet the licensing requirement of polarity items (it is not c-commanded by the negative marker). In this case, it is incorporation that is chosen:

No student read this paper

This means that even in a language where marking by polarity item is the default option, incorporation will emerge as the natural solution for marked cases. It is tempting to consider the varieties of American Black English studied by Labov (1972) as a generalised use of what is the marked option in Standard English.

But incorporation has an obvious drawback: if all negative variable markers include a negative meaning, negation being a potentially recursive operator, there is a tension between a recursive interpretation, and an interpretation as a mere negated variable marking. Applying compositionality in the normal way means that variable marking by incorporation should end up in ambiguous association: expression of a negation/marking of a variable in the scope of a negation. In a sense, this would disqualify incorporation as a robust way of unambiguously marking a negated variable. Maybe things should not be read in this way. If negative incorporation is such a widespread phenomenon in natural languages, it is probably because the recursive interpretation of negation in a clausal domain is prohibited by some specific constraint.

5. Recursivity of negation within the predicate argument domain: Principle 2

Observe that it is rather difficult to find a simple clause interpretation using more than one negation in its representation. When this actually happens, some special conditions are in general required. It has been argued in Corblin (1994, 1996) and Corblin and Derzhanski (1997) on the basis of examples from French and Bulgarian that the logical translation of a clause cannot contain more than one negation embedded under another negation. The double negation reading itself requires very special conditions to arise, which leads one to think that 'one negation per clause' is the default case. These observations seem to apply to many
languages. Whether or not they turn out to be universal, it is worth trying to put them down in
the form of a hypothesis that captures them, which is falsifiable.

Note that this tentative generalization is independent of any assumption regarding the
interpretation of any specific lexical item, and is to a large extent theory-independent. It looks
like a general constraint on the syntax-semantics interface: in the derivation of a clause, a
negation nested inside another negation is really exceptional, and one level of recursion
represents the maximal complexity allowed for the semantic representation.

We will try now to make this observation more precise. From the discussion of French
data in Corblin (1994, 1996) we see that:

1) It is rather difficult to get a translation like \(\neg \exists x \neg \exists y \ (V x, y)\), where \(x\) and \(y\) are
arguments of \(V\). Note that this seems really to be a limitation on the processing of a single
verb, since a double negation reading is unambiguously imposed by the use of two different
verbs.

\[
\text{(21) Personne ne croit personne} \quad \text{(DN reading unlikely)}
\]

\[
\text{Nobody-believes-nobody}
\]

The preferred interpretation of (21) is a mono-negative reading: \(\neg \exists x, y \ (\text{believe } x, y)\). The
truly recursive reading \(\neg \exists x \neg \exists y \ (\text{believe } x, y)\), if not impossible, requires a specific
intonation or context.

\[
\text{(22) Il n'y a personne qui ne croit personne} \quad \text{(only DN reading)}
\]

\[
\text{There is nobody who believes nobody.}
\]

The only interpretation of (22) is: \(\neg \exists x \neg \exists y \ (\text{believe } x, y)\). This shows that there is a
constraint prohibiting the truly recursive interpretation which operates within the clause
boundaries.

2) It is impossible to get a translation like: \(\neg \exists x \neg \exists y \neg \exists z \ (V x, y, z)\), where \(x\), \(y\) and \(z\) are
arguments of \(V\). Consider tell \((x, y, z)\); for the sake of simplicity we eliminate time and space
information, and a simple clause with the French verb translated as tell:

\[
\text{(23) Personne ne dit rien à personne}
\]

\[
\text{Nobody tells nothing to nobody}
\]
It has only two readings:

i) mono-negative = there is no x, y, z such that x tells y to z (default reading, a representation with a single negation).

ii) di-negative = there is no x such that there is no thing y and no person z such that x tells y to z. (Everyone tells something to someone, a representation with one negation nested inside another negation).

As suggested in Corblin (1996), the constraint relates to complex negation auto-embedding. Following Corblin’s terminology, we call negation auto-embedding a negation having immediate scope over another negation; this auto-embedding is complex if it is not eliminable "on-line" by the double negation law). If sentence (23) is pronounced without any special stress or pause, the interpretation suggested at first by almost all speakers is (i). Then some people sometimes add (especially when asked) "Well, this sentence can also mean [(ii)]". The accessibility of interpretation (ii) for (24) is highly enhanced by a marked intonation paired with a break which results in the processing of the sentence being split "into two parts":

(24) PERSONNE// ne dit rien à personne
(25) Personne ne dit rien // à PERSONNE

In question/answer pairs like (26), the only available interpretation of the answer is the double negation reading (ii).

(26) Q: Qui n'a rien dit à personne?
    A: Personne

There are in French some very rare cases making obligatory the double negation reading; all of them involve pas:

(27) Personne n'est pas venu
    Nobody-did-not come

Sentence (27) has a double negation reading only, with a scope hierarchy which mirrors left-right order: \( \neg \exists x \neg \forall x \) = everyone came. We will leave this special and
rather rare case of "obligatory double negation reading" within the clause boundaries for further investigation.

We turn now to an alleged counter example to the di-negative reading constraint. Richter et Sailer (to appear) argue that the following examples exhibit a tri-negative reading:

(28) a. Jean n’a rien dit à personne
    b. Mais non, Jean n’a pas rien dit à personne (Richter and Sailer, example (54))
    (John did not say anything to anybody/ No, on the contrary it is not the case that he did not say anything to anybody)

Richter & Sailer correctly observe that the interpretation of (28b) is:

(29) \( \neg \neg \exists y \neg \exists z \text{TELL} (\text{HE}, y, z) \)

But, contrary to what they suggest, this is not really a tri-negative reading. In fact, the relevant notion is "complex auto-embedding". (29) begins with an eliminable double negation: \( \neg \neg \exists x \neg \exists y \text{TELL} (x, y, z) \) can be rewritten as: \( \exists x \neg \exists y \text{TELL} (x, y, z) \), which is mono-negative.

A real counterexample to our claim would be a configuration like:

(30) \( \neg \exists x \neg \exists y \neg \exists y \text{TELL} (x, y, z) \)

Suppose we can find a language in which this configuration can be obtained as a translation of a single clause (with only one finite verb translating TELL). Then it would be a direct counterexample to our claim as stated in (31):

(31) CLAIM: The processing of a single clause (verb + its obligatory arguments) can only give rise to a semantic representation including at most one complex auto-embedding of negations.

Corblin (1996) and Corblin and Derzhanski (1997) argue that this constraint looks like a competence limitation on the processing of a clausal verb-argument domain. More should probably be said on this. But we leave the matter here.

The conclusion of this section is the following. If one looks at simple clauses in numerous languages, without any specific theory regarding the matters under discussion
Let us try now to recall the principles and see what we can infer from their combination:

Principle 1:

The unambiguous identification of any variable for an obligatory argument interpreted in the scope of a clause-mate negation is a meaningful and useful device. Many natural languages make this option available.

To check whether this option is available in a given language, we need to do the following test: Is it possible to express unambiguously that any obligatory argument is in the scope of the clause-mate negation in the intended interpretation of the clause?

Principle 2:

One negation per clause is the default case. One auto-embedding of negation is the highest level of complexity admissible for the processing of a single clause (i.e. for the semantic space organised by the verb-argument relation).

To check whether this is true for a given language, we need to try to get more negations in the representation of a single clause, allowing any lexical selection, and any intonational pattern. The only restriction is that it must be a single clause (containing one finite verb).

If these two principles are valid, then it seems that incorporation (morphological/or semantic) is a good solution, and probably the best one. The only drawback of incorporation would be a recursive computation of negative meanings in case of reiterated marking. However, principle 2 means that the risk is negligible, except in marked cases. The result is that many natural languages seem to use negative incorporation very liberally, just to mark lexically the fact that a variable is in the scope of a negation. Principle 1 tries to explain why negative variables should be lexically identified as such. Principle 2 tries to explain why incorporation of the negative marker (or of the negative meaning) is the most natural way to do it.

Now, as Ladusaw observes, an interesting point about the expression of negation in a clause is that there are “configurational” constraints on the distribution of negatives. Our
assumption, following a well-established idea in the literature, is that a large number of languages, including most Romance languages, obey some version of the “Neg-first” principle.


So far, we have detailed two principles that offer a suitable way of modelling the phenomenon of negative concord in its main features across languages. It was said, however, that these principles might not be enough for a fine-grained description of its manifestations in given languages. Many languages exhibit specific constraints on the distribution of negatives within the clause, many of them involving relative order between verb and negatives. This is particularly true for the Romance languages. In this section we consider this aspect and try to pin it down by elaborating on the idea of "neg-first" found in the literature. In Jespersen (1917) we find the following statement about universal tendencies in the positioning of negative expressions:

“...there is a natural tendency, also for the sake of clearness, to place the negative first, or at any rate as soon as possible, very often immediately before the particular word to be negatived (generally the verb…)” (1917:5).

The hypothesis contained in this quotation is composed of two parts, as noted by Dahl (1979). The first part is about the positioning of negatives relative to the rest of the linguistic material in a sentence, namely an early positioning. The second is about absolute positioning, and it is said that the most common position for the negative is immediately preceding the verb. Dahl shows that Jespersen’s hypothesis holds in part, with the proviso that we must consider the ‘finite element’ rather than the verb as the element that is relevant for the positioning of the negative. The finite element is the element that bears the marks for tense and mood, and possibly agreement. Dahl’s typological study shows that, for languages realising negation as a particle, the most common position for the negative is indeed the one before, and as close as possible to, the finite element, with possible variation in verb final languages (Payne 1985). On the other hand, the first part of Jespersen’s hypothesis finds little support as an independent hypothesis. In more general terms, the second part of Jespersen’s hypothesis can be taken to capture the preference for a negative to precede its focus. This is the interpretation
that we also find in Horn (1989), from whom we borrow the term of “Neg-first”. As we have seen above, the negative applies to the predicate-argument structure that represents the eventuality described in the sentence. It is in this sense that we can say that it focusses on the verb.

In this last part of our study, we work out a hypothesis on the positioning of negation in terms that are relevant for the analysis of Romance languages. We call it the Neg-first Principle. This principle is concerned with the surface order of the elements, and does not require any special assumption with respect to syntactic movement and basic positioning.

**Neg-first Principle** (general version)

A negative tensed clause must contain a preverbal negative

Recall that negatives are either the negative marker---defined above as an element that can set to false the truth value of a proposition without affecting any particular argumental position or adjunct---and n-words, i.e. negative variable markers incorporating the negation either morphologically, e.g. the Italian *nessuno*, or semantically, e.g. the French *personne*. This principle says that a simple clause must contain an element identified as ‘negative’ in preverbal position, in order to be interpreted as negated. It follows, for instance, that if an n-word occurs in post-verbal position without a negative occurring pre-verbally, Neg-first is not satisfied and the sentence is not acceptable in languages obeying Neg-first.

Let us illustrate briefly the general functioning of this principle by using Italian data (Horn himself suggests the use of some variant of the Neg-first principle to account for the Italian distribution of *non*). In Italian, *non* is the negative marker, which is realised in preverbal position, and *nessuno* is a negative variable, a case of the incorporation of negation. The meaning of negation is that the clause-mate verb must be represented in the scope of a negation, and as for the variable marking, the corresponding variable is in the scope of the clausemate negated verb.

Applying the general version of Neg-first to the relevant data means that we must find at least one n-word, like *nessuno*, or the negative marker *non*, before the verb for the sentence to be acceptable. Predictions go as follows:

(32) *X V nessuno*  
    nessuno V X  
    Nessuno V  

    if X does not contain a negative word.  
    is OK  
    is OK
However we must tune Neg-first for different languages, in order to obtain a broader coverage of the linguistic data. For instance, the general version is unable to predict that the combination in (33) is not an instance of negative concord in Italian.

(33) *Nessuno non V

Therefore, this general version of the principle will be made more precise according to the peculiarities of the languages under examination. We will discuss three cases, namely Rumanian, French and Italian.

6.1 The case of Romanian

6.1.1 Negative Concord

In Romanian the negative marker is *nu*. The Neg-first principle for this language is instantiated as follow.

**Neg-first Principle** (Romanian)

A negative tensed clause must contain the preverbal negative marker *nu.*

As it appears from the data below, a sentence containing the negative marker *nu* is interpreted as being negative. Furthermore, all negative sentences must contain *nu.*

Let us start from a simple case. Sentence (34) contains a single negation and says that there was no coming of Petre.

(34) Petre nu a venit

Peter didn’t come

In (35) there are two negative elements. The negation of the ‘eating’ event is expressed by the negative marker *nu* that, by its presence, satisfies the Neg-first principle. The n-word *nimic* lexically marks a variable in the scope of the negation.
More precisely, this sentence says that there was no eating by Petre of any sort of object. The fact that the object is expressed as a variable in the scope of a negation provides information on the events of eating by Petre that must be excluded from the set of possible models. Sentences (36) and (37) are ungrammatical because Neg-first, in the specific version for Romanian is not satisfied.

(36) *Nimeni a venit
(37) *Petre a mâncat nimic

 Romanian n-words originate from incorporation of morphological negation, e.g. nimeni (nobody), nîmic (nothing), niciodatǎ (never), nicǎierì (nowhere), etc. However, the negative element they contain is not able to satisfy Neg-first, as shown in (36) and (37). Neg-first is satisfied in Romanian by the very element nu.

(38) Nimeni nu a venit
Nobody came

In sentence (39) the n-word in subject position marks it as containing a variable in the scope of negation, whereas nu satisfies the Neg-first principle.

Finally, we can consider the case where there is more than one variable in the scope of negation.

(39) Nimeni nu a zis nimic
Nobody said anything

Here again each n-word in argumental position is interpreted as a variable in the scope of negation. The sentence says that there was no event of saying having someone as the agent and something as an object. The Neg-first Principle is satisfied through the marker nu.

6.1.2 Double negation
The Neg-first principle, in its instantiation for Romanian, allows us to express an interesting observation. The negative marker *nu*, used to mark clausal negation via the satisfaction of the Neg-first principle, doesn’t give rise to a logical double negation reading. From this observation we draw the following tentative empirical generalisation:

(Tentative generalisation)

Negative markers that satisfy Neg-first do not give rise to DN readings.

In Romanian, it seems that it will not be possible to get a logical double negation reading in a simple clause containing *nu*. In the few cases where it may be claimed that there is a DN reading, it is not available for all speakers and it is perceived as being paired with special conditions. For instance, (40) has a NC reading for all speakers, and this is the preferred interpretation.

(40) Nimeni nu iubeste pe nimeni
    Nobody does not love anybody

This sentence is reported to have a DN reading only by some speakers, provided it is pronounced with a particular contour, and that those very words (*love* and *nobody*) are used. Note that, in any case, if a DN reading occurs it can only be supported by two n-words.

6.2 The French case

6.2.1 Negative concord

The second case we are going to consider is French. As will immediately become clear, in order to appreciate the manifestations of NC in French, one has to consider the development of this language over time. Its behaviour with respect to the principle has not been stable, but we may understand its coherence by taking a diachronic view. The French negative marker is *ne*. The Neg-first principle for French is instantiated as follows:

**Neg-first Principle** (French, first version)
A negative tensed clause must contain the preverbal negative marker *ne*.

Old French appears to obey the Neg-first principle.

(41) *Jeo ne di*

I not say

(41) shows that *ne* was a full expressor of negation in Old French. Then the language goes through what is known as Jespersen’s Cycle. The negative marker *ne* is strengthened by a postverbal minimiser (Bolinger 1972). Initially, there is a selection of candidates for this function, e.g. *pas, point, mie, goutte*. But then in modern French this selection of elements is cut down to the first one, and possibly the second one in fairly high registers. The fact that a given element specialises for the function of negation strenghtener is accompanied by its promotion to the status of element fully able to express negation. This phenomenon can be analysed in the following way. The choice is reduced to a single item, and at the same time this item loses its function of identifying/expressing the minimal unit on a particular scale. Its lexical content, making the identification of a scale possible, appears to be lost. The inferential effect of crossing out all cases on a scale by crossing out the strongest position is generalised from scales of a given type to any object. The result is the interpretation of the minimiser *pas* as a negative marker. This is the situation we find in standard modern French, where

*ne V pas*

is the unmarked way of expressing negation in a sentence. The occurrence of *ne* satisfies the Neg-first principle, but it is not enough to express the negation. This task is performed by *pas*. The revised version of the principle presented below makes the change explicit by taking off the characterisation of *ne* as negative:

**Neg-first Principle** (French, second version)

A negative tensed clause must contain the preverbal marker *ne*.

This is the version that is satisfied in standard Modern French.
(42) Jean ne mange pas  
John is not eating

In (42) the effect of negating the existence of an ‘eating’ by Jean is expressed by _pas_, whereas _ne_ is there to satisfy the Neg-first principle.

It is worth noting that French has followed a different path with respect to Romanian in building its selection of negative expressions. In French the use of minimisers is the general strategy, adopted also for marking variables in the scope of negation. For instance, besides _pas_, which somewhat expresses the minimal quantity of a process (more generally of an eventuality), we find _personne_, used for the minimal unit for human beings, and _rien_, which is the minimal unit for things (from Latin _rem_). In short, the use of minimisers is a typological choice for French, as opposed to the possibility of incorporating negation that we have observed in Romanian.

(43) Jean ne mange rien  
John isn’t eating anything
(44) Personne ne mange  
Nobody is eating

In (43) the n-word in object position indicates that there is no eating of any object. The function of _ne_ is to comply with Neg-first. The same line of reasoning applies to (44), where it is the subject position that is marked as filled by a variable in the scope of negation. _Ne_ is there to satisfy the Neg-first principle, since the n-word _personne_ is not able to satisfy it even if it occurs in preverbal position. The same is true of (45), where we have two variables in the scope of negation.

(45) Personne ne mange rien  
nobody eats anything

Finally, in spoken/colloquial French, we observe the disappearance of _ne_ in many cases, e.g. with one postverbal n-word, as in (46), with the ’expressor’ of negation, as in (47) and with more than one n-word, as in (48).
This situation is captured in our last version of Neg-first for French. Here, the vacuity of the operation of satisfying a version of Neg-first via a marker that is not a negation is made explicit by allowing the marking to be optional.

**Neg-first Principle** (French, last version)

A negative tensed clause can contain the preverbal marker *ne*.

6.2.2 Double negation

The sentence in (45) also has a DN reading. This second reading is simply not available for (43) and (44). The absence of a DN reading in these cases can be easily explained by noting that *ne* is the element satisfying Neg-first (see what was said for Romanian). This is true for any stage of French, i.e. whether *ne* is negative (Old French) or not (modern French). This marker cannot enter DN. In contrast, the case of (27) cannot be analysed in such a straightforward way. We note that DN readings seem to be more accessible in French than in other Romance languages. Furthermore, they seem not to need a special prosodic pattern in certain cases. We have no explanation for this phenomenon as yet.

There is another fact that our analysis in its current version seems unable to predict. Sentence (49) is perceived as marginal, but if accepted, it has only a DN reading. What has been said so far cannot distinguish (49) from (45).

(49)  **Personne ne mange pas**  
      Nobody does not eat

6.3 The case of Italian
6.3.1 Negative concord

As a last case for this review, we consider Italian. This language appears to obey a modified version of the principle, insofar as the principle seems not to be tied to a particular realisation of what counts as negative marker, as it can be observed by the acceptability of (51) beside the more predictable case of (50).

(50) Non mangia
    S/he is not eating
(51) Nessuno mangia
    Nobody is eating

Sentences (50) and (51) both contain a negation and can be represented in terms of a negated predicate, with a difference in their respective argument list. The distinction is that only in (51) it is made clear that the agent position cannot be instantiated, or rather that no matter what its instantiation is, the predicate-argument structure describes an event that does not take place (when considered relatively to the instantiations of the other arguments).

At the same time, the unacceptability of sentence (52) shows that some form of Neg-first must be satisfied in Italian.

(52) *Vede nessuno
    [s/he] sees nobody

We propose to modify slightly the constraint imposed by the general version of the Neg-first principle in order to capture the Italian distribution. The type of modification we are looking for can be obtained simply by enabling any morphologically/lexically marked negative item to satisfy the constraint.

**Neg-first Principle** (version for Italian)

A negative tensed clause must contain one and only one lexically negative preverbal marker

The lexically negative markers relevant to the principle are *non* and n-words taken in the strict sense of the term, i.e. those items that are morphologically marked as negative. Examples (51)
and (53) show that an n-word such as *nessuno*, that originates from the addition of a negative on an indefinite, can satisfy the principle.

(53) Nessuno vede nessuno
    Nobody sees anybody

Broadening the set of elements that can satisfy the principle has, as a consequence, the situation whereby a sentence could contain more than one candidate to the function of ‘satisfier’ of the principle. This does not seem to be the case, at least not in a standard sentence pronounced in an unmarked manner. For this reason, the principle requires the presence of ‘one and only one’ marker in preverbal position.

The restriction on the presence of a single preverbal marker might sound too strong to speakers who find sentences such as (54) acceptable. However, the point here is precisely that this sentence is not accepted by every speaker and does not have the same status as (50), (51) or (53).

(54) ?*Nessuno niente ha fatto
    Nobody anything did

If the expression of negation via the marking of an argumental position does not satisfy Neg-first, for instance because it occurs in postverbal position, then this principle is enforced under its general version, i.e. licensing is done via the basic negative marker *non*, as exemplified in (55).

(55) Non vede nessuno
    S/he does not see anybody

Recall that Italian formed almost all its n-words by affixation of a negative marker. The fact that only morphologically negative elements can satisfy the principle shows that it would be incorrect to perceive the modification required for Italian as a case of constraint relaxation. Rather, it is a case of generalisation that can be appreciated in full by taking a diachronic view. From the early stages in the development of Italian, the presence of the negative marker on the n-word in preverbal position (through morphological incorporation) is perceived as sufficient for the realisation of negation. The situation remains stable across centuries. Posner
(1984) observes that ‘examination of popular and dialectal texts reveals no evidence of the use of [the negative marker] after [an n-word] in Central and Southern Italy or in Sardinia, and informants claim that even illiterate speakers do not make this “mistake”.’ (1984:16).

The sentence in (56) shows that the satisfaction of the Neg-first principle is possible only via a morphologically negative n-word. Mai (never) has negative meaning but originates from the positive Latin form magis ‘more’.

(56)  ?* Mai vede nessuno
       S/he never sees anybody
(57)  *Ha mai parlato
(58)  Non ha mai parlato
       S/he never spoke

Posner also observes that ‘mai did not acquire intrinsic negative character until after the fourteenth century’ (1984:13), whereas nessuno and niente were already well established in their function of full negatives in preverbal position in Dante's and Boccaccio's literary production, i.e. at the beginning of the fourteenth century. Note that apparently this is the only n-word that can cooccur with another n-word in preverbal position. There are no order constraint, compare (59) and (60) which do not exhibit relevant differences.

(59)  Nessuno mai mi ha chiesto di parlare
       Nobody ever asked me to speak
(60)  Mai nessuno mi ha chiesto di parlare

However, the issue of the behaviour of mai requires further study. The case in (61) can be reduced to (56) and is accounted for. But something will have to be added to account for the acceptability of the emphatic sentence in (62).

(61)  *Mai leggerò nessun libro
       I will never read any book
(62)  Mai leggerò questo libro
       I will never read this book
The case of emphatic/strongly affective sentences such as (63), which requires special prosodic contour, also needs further specification.

(63) (Sono tutti degli ingrati.) Nessuno non è venuto
(they all never appreciate) nobody came

6.3.2 Double negation

The violation of Neg-first in Italian leads either to a case that is not acceptable or to a case of logical double negation, since there is more than one element that is a good candidate for satisfying the principle.

(64) Nessuno non vede questo cartello
Nobody does not see this sign
(65) Mai nessuno non guarda la partita
Nobody ever doesn’t watch the match

Furthermore, on the basis of this principle it is possible to predict where the splitting into two negations is going to be located. The partition takes place with respect to the element that satisfies the principle and all the other negations coming after it are taken in a bunch, as it appears from the gloss under example (66).

(66) Nessuno non porta un fiore a nessuno
Nobody does not bring a flower to anybody.

In short, *non* introduces a double negation if it is preceded by another negative, but it does not enter in a double negation with an n-word following it—a consequence, in fact, of the role it plays in the satisfaction of the Neg-first principle.

7. Conclusion

In this paper, we have put forth a proposal for covering salient data regarding the expression of negation in some Romance languages without making use of a specific "negative concord"
property. Our assumption is that what is called "negative concord" is in fact a phenomenon which results from different constraints relative to the expression of negation in the verb-argument domain.

The key-point of our view is that the most general features of the so-called "negative concord" are the result of two forces: on the one hand there is the lexical marking of negated variables, and on the other hand a competence limitation which tends to constrain negation as a clausal feature, hence making the recursive computation of negations in the verbal domain very unlikely an option. We assume that the promotion of "incorporation"—i.e. the lexical marking of negated variables via the repetition of the negative marker itself—to the level of option widespread across natural languages is the consequence of these two tendencies.

Finally, a more ad hoc and restricted generalization has been assumed to capture a property which seems common to most Romance languages: the general idea is that a lexical item expressing negation must occur pre-verbally.
8 References


Corblin Francis. 1996. Multiple Negation Processing in Natural Language. Theoria 17, 214-259

Corblin Francis and Ivan Derzhanski. 1997. Multiple Negation, optional arguments and the reification of eventualities. in Corblin et al. (eds.) Empirical Issues in Formal Syntax and Semantics, Berne, P. Lang


Deprez Viviane. 1997. Two types of negative concord. Probus 9, 103-143


Jespersen Otto. 1917. Negation in English and other languages. reprinted in Selected Writings of Otto Jespersen, George Allen and Unwin


Ladusaw William. 1991. Interpreting negative concord structures. text of the talk delivered at LSA


Przepiorkowski Adam and Anna Kupsc. 1999. Eventuality Negation and Negative Concord in Polish and Italian. in Borsley R. and A.Przepiorkowski (eds.) Slavic in HPSG, CSLI

Richter Frank and Martin Sailer, to appear, A lexicalist collocation analysis of sentential negation and negative concord in French. in Kordoni V. (ed.) Tubingen Studies in Head-Driven Phrase Structure Grammar, CSLI

Solà Jaume. 1972. La negació. in *Estudis de sintaxi catalana*, Barcelona: Edicions 62

Francis Corblin
UFR Langue Française
U. Paris IV-Sorbonne
1, rue Victor Cousin
F - 75230 Paris Cedex 05
Francis.Corblin@paris4.sorbonne.fr

Lucia M. Tovena
UFR Angellier
U. Lille 3
BP 149
F - 59653 Villeneuve d’Ascq Cedex
tovena@univ-lille3.fr