The bi-polar any

Jacques Jayez
EMEFSS
54, Bld Raspail,
75006 Paris, France
jjayez@dial.celeste.com

Lucia Tovena
IRST
via Sommarive 18,
38050 Povo Trento, Italy
tovena@irtc.it

Salford Conference on Negation: Syntax, Semantics and Pragmatics,
Salford, 30 October - 1 November 1998

1 The problem

In the literature on any, there is a traditional distinction between Polarity
Sensitive (PS) any and Free Choice (FC) any.

PS any
ok in negative and interrogative contexts (Mary didn’t read any book; I
doubt that she read any book; Did you read any book?, etc.)
Akin to an indefinite or an existential quantifier

FC any
ok in some positive contexts and in generic contexts (Mary read any book
proposed by her advisor, Any owl hunts mice, etc.)
Akin to a universal quantifier

There have been quite a few attempts to ‘unify’ any, that is, to derive its
overall distribution while postulating a single any. Some examples,

<table>
<thead>
<tr>
<th>Existential any</th>
<th>Universal any</th>
</tr>
</thead>
</table>

These attempts are extremely difficult to assess. For instance,
the notion of ‘indefinite’ is not so clear,

Indefinite any shows strong ∀ properties, There is no general agreement on
the mechanism which would be responsible for the indefinite ∀ drift (see
Haselmath 1993 for a discussion).

the properties of any as ∀ do not square with the observed distribution,
Moreover there are strange parallelisms between any and le moindre (French)
(Tovena & Jayez 1998). However le moindre is not considered an indefinite
nor a universal quantifier in French.
In this talk, we skirt the indefinite/∀ controversy and propose an uncommitted theory.

2 Bi polar (strong) and weak any. Sketch

<table>
<thead>
<tr>
<th>PS any</th>
<th>FC any</th>
</tr>
</thead>
<tbody>
<tr>
<td>bi polar negative any</td>
<td>bi polar positive any</td>
</tr>
</tbody>
</table>

The ‘weak’ any is Horn’s (1996) indiscriminative, analogous to n’importe quel
N in French. It is to be found at least in choice imperatives (Pick any card)
and may permissions (You may pick any cake).
The bi polar or ‘strong’ any has positive and negative polarity.
Polarity means: informational emphaticness (Israel 1996).

Why ‘strong’?
Strong any has modal properties: non situationality (Eisner 1994, Dayal
1995, 1998), non individuation (Tovena 1996), arbitrariness (Kempson 1985,
Tovena & Jayez 1997b, c, 1998), variation (Tovena & Jayez 1997b, Giannakidou
1997b). These properties are manifestations of informational emphaticness
at the model theoretic (situation theoretic) level.

An open question: can these ‘manifestations’ be derived in a strict sense or
are they just ‘motivated’? An answer would require (at least) a thorough
Linguistic study. See Tovena & Jayez 1997a, Giannakidou 1997a, Lee
(Chungmin) 1997 for some facts.

Some examples to get the flavour (Lee & Horn style (1994))
Mary didn’t read any book = ‘Mary did not read (even) the book(s) she was
expected to read'.

impcative: ‘Mary read absolutely no book’.

Did Mary read any book? = ‘I wonder whether Mary read even the book she might have reasonably read’.

impcative: ‘Really, I wonder whether Mary read some book’.

Mary read any book which was on the reading list = ‘Mary read (even) the book(s) on the r.l, she was not expected to read’.

impcative: ‘Mary read absolutely every book on the reading list’.

See Horn 1972, Fauconnier 1978, Ducrot 1980, Kay 1990, Kadjou & Landman 1993, Lee & Horn 1994. Horn 1996, etc. All these people (the ‘scalarists’) draw our attention to a very important idea,

Scalar implicatures are not pragmatic sugar. They form the semantic core of some lexical items. Polarity is the way in which scalar implicatures get lexicalized.

What about the ‘modalists’ (Eisner, Dayal, Giannakidou, the present authors)? They draw our attention to the fact that any hates plain episodic descriptive information in some cases: Mary read any book. They relate this fact to some modal strength conveyed by any.

Each party has explanations of its own, but we propose rather that the two cooperate.

3 The weak any

Just mentioned in passing in this talk, it deserves a study of its own. Some brief remarks.

(Parly?) parallel with n’importe quel N (lit. ‘not matter what’ N).

1 a. Prends n’importe quelle carte
(Pick any card)

b. Tu peux prendre n’importe quel gâteau
(You may pick any cake)

The crucial point is that neither any nor n’importe quel are modally strong in (1): they do not mean ‘any card in any possible situation’ or ‘any cake in any possible situation’ (= Dayal’s analysis in her 1998 paper).

More debatable: is there a parallelism in (2)?

2 a. N’importe quel étudiant sait ça  
(Any student knows that)

b. Est-ce que Marie a lu n’importe lequel de ces livres?
(Did Mary read any of these books?)

c. A la dernière course scolaire, Sandra a couru plus vite que n’importe quelle autre fille de sa classe
(At the last college race, Sandra ran faster than any other girl in her class)

For a general approach to FC items, see Giannakidou 1997a,b.

4 A scalar tale

We rephrase the affinity between any and scalar implicatures in an information based terminology (see Krift 1995, Israel 1996 for a general introduction to the connection).

4.1 Definitions

Our basic constructs are total orders $S = (\Delta, \leq)$, where $\Delta$ is a set of degrees. N.B. We assume that $\leq$ is total to simplify the presentation. Nothing essential hinges on this choice.

Let $P$ be a gradable property of form $\lambda x, \delta, \phi(x, \delta)$, where $\phi(x, \delta)$ means that $x$ satisfies $\phi$ at some degree $\delta$. We note $\phi_{\delta}$ the property of satisfying $\phi$ at a degree $\delta$, i.e. $\phi_{\delta} = \lambda x, \phi(x, \delta)$.

3 Expectation ordering

An expectation ordering (e.o.) on an order $S$ w.r.t. properties $P$ and $Q$, is an order $S_{P,Q} = (\Delta_{S_P}, \leq_S)$ such that $\Delta_{S_P} = \{ [P_{\delta}, Q] : \delta \in \Delta_S \}$.

Intuitively, an e.o. is an ordering of how likely it is for an object $o$ to be a $Q$ object if it is a $P$ object at some given degree.

$[P_{\delta}, Q] \leq_S [P_{\delta'}, Q]$ means that an object $o$ which has the property $P$ (a $P$ object) at a degree $\delta$ is not likelier to be a $Q$ object than an object $o'$ which has $P$ at a degree $\delta'$. 

The bi-polar any - J. Jayez & L.M. Touéna
Example,

\[ P = \lambda x, \delta. \text{interesting book}(x, \delta), \quad Q = \lambda x, \exists y(\text{read}(y, x)) \]

Suppose we have an isomorphism

\[ \langle P_b, Q \rangle \leq_{S'} \langle P_{b'}, Q \rangle \text{ iff } \delta_b \leq \delta_{b'} \]

For books \( b, b' \), if \( b \) is not more interesting than \( b' \), then \( b \) is not likelier to be read than \( b' \).

(4) **Scalar domain**

Let \( S'_{S,P,Q} \) be an e.o. The scalar domain (s domain) in \( S' \) of a property \( P_b \) is the set \( \{ \delta_j : \langle P_j, Q \rangle \geq_{S'} \langle P_b, Q \rangle \} \).

Intuitively, a s domain is the set of degrees which plausibly ‘trigger’ \( Q \) given that the degree \( \delta_i \) of \( P \) has triggered \( Q \). It’s a sort of a fortiori classification: if you have \( Q \) when you reach a \( \delta_i \) of \( P \) ness, then you have also \( Q \) when your reach a greater degree of \( P \) ness.

A proposition is s relevant whenever it has a maximal s domain.

(5) s relevance

Let \( S'_{S,P,Q} \) be an e.o. \( P_b \) is s relevant in \( S' \) when its s domain is \( \Delta \) itself.

Associated with predicates and dimensions, there are normal e.o. For instance, interesting/relevant/epochal books are more likely to be read than dull/relevant/insignificant ones, etc.

(6) **Informational relevance for strong any**

\( \phi(\text{any N}) \) is felicitous w.r.t. some gradable property \( P \) iff

a. \( P \) entails \( N \),

b. for some \( \delta_i \), \( P_b \) is s relevant in some normal \( S'_{S,P,\phi} \).

(7) \forall value of strong any (first version)

If \( \phi(\text{any N}) \) is felicitous w.r.t. some property \( P \), it triggers the usual scalar conventional implicature \( \forall x(N(x) \Rightarrow \phi(x)) \).

### 4.2 Examples

(8) a. Mary read any book which was on the reading list

b. Mary didn’t read any book

c. Any student knows that

d. At most two students read any book

e. Did Mary read any book?

\[(8a)\]

Let \( P = \lambda x, y, \text{book}(x) \land (\text{book}(y) \Rightarrow \text{likely to be read}(x,y)) \) and \( Q = \lambda x, \text{Mary read}(x) \).

\( P \) selects all pairs \((x,y)\) such that \( x \) is a book on the r.l, and \( y \) is the plausibility of \( x \) being read if \( x \) is a book on the r.l. \( P \) obviously entails the property of being a book on the r.l, that is, \( P \) entails \( N \).

Let \( \delta \) be min \( \Delta \) and \( S' \) be the normal e.o. defined by the isomorphism above

\[ \delta_i \leq \delta_j \text{ iff } \langle P_{b'}, Q \rangle \leq_{S'} \langle P_{b}, Q \rangle \]

Then any is appropriate w.r.t. \( P \), \( Q \) and \( S' \), The scalar domain of \( P(\delta) \) is, by def. (4), \( \{ \delta_j : \langle P_{b}, Q \rangle \geq_{S'} \langle P_{b}, Q \rangle \} \). By the postulated isomorphism, this set is the set of degrees, \( \Delta \) itself. So \( P(\delta) \) is s relevant by def. (5) and any is appropriate by def. (6).

\[(8b)\]

Let \( P = \lambda x, y, \text{book}(x) \land (\text{book}(y) \Rightarrow \text{likely to be read}(x,y)) \) and \( Q = \lambda x, \text{Mary didn’t read}(x) \).

The reasoning is analogous with \( \delta_i = \max \Delta \) and the reverse isomorphism this time.

\[(8c)\]

Let \( P = \lambda x, y, \text{student}(x) \land (\text{student}(y) \Rightarrow \text{likely to know that}(x,y)) \) and \( Q = \lambda x, \text{knows that} \).

Let \( \delta \) be min \( \Delta \) and an isomorphism as in (8a). The reasoning is similar to the (8a) case.

\[(8d)\]

Let \( P = \lambda x, y, \text{book}(x) \land (\text{book}(y) \Rightarrow \text{likely to be read}(x,y)) \) and \( Q = \lambda x, \text{at most 2 students read}(x) \).

N.B. There are two residual problems with this (kind of) example. First, unexpectedly, le moins is not felicitous in the corresponding French examples: *Deux étudiants au plus ont lu le moindre livre.* Examples such as *Il n’y a que deux étudiants qui aient lu le moindre livre* ("There are only...") are ok, however. Second, there is an additional scope constraint since (8d) entails but is not equivalent to \( \forall x(\text{book}(x) \Rightarrow \text{at most 2 students read}(x)) \).
Contrary to Dayal (1998), we need not assume that may permissions and choice imperatives have modal properties, since we ascribe them to weak any. So we avoid the counter-intuitive assumption that You may pick any card—Pick any card might at some level refer to any card in the world. This move is corroborated by the difference in French between the strong tout and the weak n’importe quel.

(11) a. Tu peux prendre toute carte
    (Really? Any card I find in the world?)

    b. Tu peux prendre n’importe quelle carte (≈ (1-b))
    (From the pack, of course)

5.2 Subtrigging

Dayal (1995, 1998) emphasizes the importance of subtrigging, i.e., the fact that certain types of postnominal modifiers (for instance, relative clauses and postnominal adjectives) redeem any sentences which would be anomalous otherwise. Subtrigging is illustrated by the contrast in (12).

(12) a. Mary read "any book (= (9))

    b. Mary read any book which was on the reading list (=8-a))

    c. Mary read any book proposed by her advisor
    d. Mary checked any result which depends on Craig’s theorem

    e. Mary touched "any book which was on the shelf

The most obvious effect of subtrigging is to restrict the domain. In (12-b), the infelicitous implicature that ‘all’ books have been read is replaced by the more natural implicature that all books on the r,l have been read.

However, not just any restriction redeems the any phrase, since ‘accidental’ modifications of the N, as in (12-e), do not improve the sentence. What happens in (12-e) is that no clear connection between being touched by Mary and being on the shelf emerges. Accidental in φ(any N) = contingent connection between the N property and φ.
5.2.1 Modal force and arbitrariness

Any forces an unrestricted ∨ reading. In (12-a), the fact that Mary read every book is not presented as a simple enumeration of facts (Mary read the book A and the book B and . . .), because simple enumerations ignore the scalar hierarchy.

Is emphaticness compatible with contingency? Not really. Let’s look at (12-b) again. If Mary read absolutely every book on the r.l., she must have had some reason/disposition/obligation for/of doing that. Problem: why can’t we understand (12-b) as meaning ‘it turned out (by luck) that Mary read every book on the r.l.’?

We consider the notion of arbitrariness (Fine 1985,1988,1989, Meyer Viol 1985). Arbitrary objects = objects which possess all and only the properties that define their object class. They are deprived of all individual properties (Plato’s Ideas are not very different), (12-b) = ‘Mary read even the book(s) on the r.l. that she was not likely to read’.

Thus, the individual (negative) characteristics of books on the r.l. did not impinge on Mary’s reading.

Implicature of arbitrariness: so, it’s in virtue of being books on the r.l. that books are read by Mary, not in virtue of being this or that book. Otherwise, some of the books (the least interesting, relevant, etc.) wouldn’t have been read.

So, there is an essential connection between being a book on the r.l. and being read by Mary (Fine again, 1995, on the logic of essence).

So, accidental modifiers are not welcome: they conflict with the natural implicature of arbitrariness.

An open question: is this implicature ‘active’ or lexicalized (or both)? It is difficult to say. It seems that the same implicature is available independently from the use of any.

(10) Mary read even the most difficult books

(10) strongly suggests that Mary didn’t just happen to read difficult books by luck (or misfortune).

Technically, ‘arbitrarification’ consists of making the set of properties of individuals vary (one then gets ‘slim’ or ‘fat’ individuals), and show that putting individuals ‘on a diet’ does not change anything essential as long as the slim individuals keep some properties which are essential to them (Fine 1988). Interestingly, ∨ has then a modal force.

5.2.2 What is subtrigging?

The nature of subtrigging remains somewhat mysterious. In some cases, the subtrigger points to a particular set (the books on the r.l.), But in (12-d), this is far from clear. The sentence should mean ‘Mary checked any result in the world which depends on Craig’s theorem’. However, we are not committed to this interpretation. If Mary works on a mathematical project which produced a particular set of results, (12-d) is fine (see Tovena & Jayez 1998 for some reservation on Dayal’s (1998) approach from this point of view).

In (12-d), we may assume that the dependency is between the property of being a result (from a fixed set) checked by Mary and the property of depending on Craig’s theorem. In general, in a form φ(any N wh. S), we may assume if needed that the dependency is between φ ing N and (wh. S) ing. Now, the scalar movement does not take us too far because the implicature is on the minimal/maximal element which Ss inside a fixed set. This suggests that we modify the definition (7).

(7) bis ∀ value of strong any

If φ(any N) is felicitous w.r.t. some property P, it triggers the usual scalar conventional implicature ∀x(ψ(N(x) ⇒ φ(x))) for ψ(N) = N or N\textsuperscript{fixed} ∧ restr(N\textsuperscript{0}), where N\textsuperscript{fixed} is the characteristic predicated of a fixed set of N\textsuperscript{0} objects and restr(N\textsuperscript{0}) that of the restriction (subtrigger) in N.

5.2.3 Negation and Non Locality

At first sight, there are no problem with (8-b).

(8-b) Mary didn’t read any book

If we consider a fixed domain of books, the strong scalar implicature (that Mary didn’t read any book at all in the world) entails that Mary didn’t read the books in the fixed domain. But,

1. The meaning assigned to (8-b) is not the right one if the speaker is supposed to refer to a fixed set of books. We have the same problem as
Dayal with Pick any card and You may pick any cake.
2. The meaning assigned to (13) may perfectly be 'accidental'. If Mary is absent, she didn’t touch any object on her desk, but she had possibly no special intention or disposition. Arbitrariness is not required with negation.

(13) Mary didn’t touch any object on her desk

However, there is a weaker notion which includes arbitrariness as a special case,

Non Local (NL)

A proposition is non local when its truth in a situation is ‘global’, that is, not dependent on any particular spatio temporal region of the situation, NL is best expressed in situation semantics (Barwise & Perry 1983, Devlin 1991) where the notion of region is built in the representation.

Intuitively:
(13) respects NL because there is no set of events which can prove this proposition. One may not select a set of spatio temporal regions in the situation where (13) is true, and conclude that Mary didn’t touch any object on her desk by inspecting those regions. One has to select ALL the regions, or a set of regions which cover the WHOLE situation, that is, one has to consider the whole situation.

(12-b) respects NL because the dependency it expresses may not be proven by selecting a set of spatio temporal regions. A dependency always presuppose some logical ‘glue’ to connect causal factors and observed effects, and, in a sense, logics is everywhere.

NL in a simplified version of situation semantics

All we need is the notions of infon: propositional objects which are spatio temporally situated, p, q, etc, denote infons. The spatio temporal location of an infon represents the spatio temporal interval at which its propositional content is true.

For instance the infon ‘Mary read Vinceland at 3 in the garden’ is situated in the garden at 3. We note r(p) the spatio temporal interval at which an infon holds.

A situation s is described by a collection of infons Σ. When an infon p is true in a situation s, s is said to support p (written as s |= p). Infons can be atomic or complex. Usually, complex infons are obtained from quantification and boolean combination, Σa is the set of atomic infons supported by s.

Infons which pertain to the same spatio temporal region define a regional view. Let r (a region) be a spatio temporal interval on a situation s.

(14) Regional view

Let s be a situation. A regional view Σr(s) on s is the set of atomic infons defined by: p ∈ Σr(s) iff p ∈ Σa, and r(p) ⊆ r.

(15) Omitting regions

The situation s omits a region r whenever it supports no infon at r, or, equivalently, its regional view Σr(s) is empty (the situation is ‘blind’ at r).

Now, we make situations ‘shrink’, that is, we suppress some regions.

(16) Shrinking

Let s be a situation. A shrinking of s is any situation s’ such that:
1. s’ omits a region r that s does not omit.
2. s’ is like s for all regions that s does not omit, i.e., if s does not omit r, Σr(s’) = Σr(s).

Finally, we say that a proposition is non local in s if it may be proved in any shrinking of s.

(17) Non locality

Let s be a situation. An infon p is non local in s iff no shrinking of s supports p.

For (13), suppose that the region of Mary didn’t touch any object on her desk is r. Consider the situation at r (the set of ‘facts’ at r). Make it shrink. How do we know that, in the subregions we erase, Mary did not touch an object on her desk? We don’t, therefore by shrinking we destroy the possibility of proving (13), or, alternatively, (13) does not depend on a set, be it infinite of particular regions of the initial situation. Note the difference with Nelson’s (1946) strong negation, which is precisely, not situational.

For (12-b), suppose that we make the initial situation shrink. We might still keep the episodes in which Mary read the books from the r.l. So, prima facie, we might prove (12-b). But we have also the dependency between properties. This information is not local. We may not prove a dependency if we do not assume logical principles/laws, which are ‘everywhere’ (Russellian in the situation semantics/jargon). The dependency evades the initial situation itself.
because it resorts to principles valid in all situations. Therefore, shrinking may not affect the non locality of a dependency-based judgment.

Finally, we impose a condition which purports to capture the modal strength common to positive and negative polar any.

(18) Non locality of any

A sentence \( \phi(\text{any} N) \) is felicitous only if it expresses a non-local infon.

Note that, in the case of Mary read \( \text{any}\) book, either the sentence refers to a fixed set of books and non locality is violated or it refers to all possible books and it is strange for that reason.

6 Conclusion

The distribution of any may be best understood by replacing the traditional NPI and FC partition with a classification that groups together positive and negative polarity (strong) manifestations while keeping apart indiscriminative (weak) uses.

The talk has concentrated on strong any and has shown that analyses based on scalar and on modal properties of any both support this new perception of the polarity sensitivity of the item.

References


Lee, Young-Suk & Laurence R. Horn, 1994, Any as Indefinite plus Even, ms. Yale University,
Meyer Viol, Wilfried P.M., 1995, Instantial Logics, An Investigation into Reasoning with Instances, ILLC dissertation, Amsterdam,