Gennaro Chierchia
University of Milan Bicocca/Harvard University
July 15 2005

Aspects of recursive pragmatics: Managing multiple alternative sets.

(1) Pragmatics and grammar.

a. What is the relation of grammar to speaker/hearer's driven enrichments of meaning (rooted in Gricean principles)?
b. How are scalar implicatures (SIs) computed?
c. How are other lexically induced implicatures (e.g. those associated with polarity phenomena) computed?
d. "Globalism" vs. "Localism" vs. "Representationalism"
   i. Globalism: implicatures are computed at the level of root sentences (Sauerland, Spector)
   ii. Localism: implicatures are computed locally/recursively (Chierchia)
   iii. Representationalism: Implicatures are driven by syntactically (LF) projected operators (Fox).

(2) Polarity sensitivity of SIs

a. We'll hire (either) Bill or Sue (but not both)
b. If we get enough money, we'll hire (either) Bill or Sue (but not both)
c. If we hire (either) Bill or Sue, we'll finally reach a good critical mass
c'. * If we hire Bill or Sue but not both, we'll finally reach a good critical mass
d. If we hire either Bill or Sue, who are senior, there will still be money left for a junior appointment.
d'. If we hire either Bill or Sue but not both,….
e. Many of them complained (though not all)
f. If the assignment was too hard, many of them complained (though not all)
g. If many of them complained (*though not all), we are in trouble
h. …though it is not the case that if a few of them complained we are in trouble
i. If many of them complained, we are better off than if all of them did.

Similarly for other scales.

j. < or, and>
k. <some, a few, many, most, every>
l. <not all, few, no>
m. <possible, necessary>
n. <one, two, three,….>

(3) Asymmetry in the distribution of enriched readings.

a. Canonical enrichment: many→ many and not all, or → or but not both, etc.
b. Canonical enrichments are dispreferred in DE contexts (where they are replaced by a "less local" kind of implicature – cf. (2h)).

(4) The Gricean take.

a. some of your students complained
b. many of your students complained
c. all of your students complained

i. The speaker chose to utter (b) over (a) or (c), which would have been also relevant
ii. (c) entails (b), which entails (a) [the quantifiers form a scale]
iii. Given that (c) is stronger than (b), if the speaker had the info that (c) holds, she would have said so [quantity]

iv. The speaker has no evidence that (c) holds
v. The speaker is well informed on the relevant facts
Therefore:
vi. The speaker has evidence that is not the case that (c) holds.
[(vi) – the epistemic step – doesn't really follow from Gricean maxims]

(5) Alternatives and operators (I)

a. $\|$ many of them complained $\|^{ALT}$ =
{ $some_D$(of them)(complained),
$many_D$(of them)(complained),
$every_D$ (of them)(complained)}

b. O_C(many of them complained) = $many_D$(of them)(complained) $\wedge$
$\forall p\ p \in C$ [p $\rightarrow$ $many_D$(of your students)(complained) $\subseteq$ p}

c. $many_D$(of your students)(complained) $\wedge \neg$ all_D (of your students)(complained)
[Krifka, van Roy, Fox; strong similarity with focus semantics]

(6) Local and less local applications of O.

a. If many of them complained we are in trouble

b. O(If many of them complained, we are in trouble) =
$many_D$(of them)(complained) $\rightarrow$ we are in trouble
$\wedge \neg$ (a few_D(of them)(complained) $\rightarrow$ we are in trouble)

c. $\Downarrow$ {
…
a few_D(of them)(complained) $\rightarrow$ we are in trouble
$many_D$(of them)(complained) $\rightarrow$ we are in trouble
$ever_D$(of them)(complained) $\rightarrow$ we are in trouble
}

d. If many of them complained, we are better off than if all did

e. O ($many_D$(of them)(complained))$\rightarrow$ we are better off than if all did =
$many_D$(of them)(complained $\wedge \neg$ all_D(of them)(complained) $\rightarrow$
we are better off than if all did

(7) The asymmetry problem.

a. Insert O freely. Everything else being equal, go for the strongest interpretation.

b. If John has two cars, he'll either buy a boat or a motorcycle

i. (Op$\rightarrow$Oq) ii. O(p$\rightarrow$q)
$\Downarrow$ $\Downarrow$

c. iii. (Op $\rightarrow$ q) $\Leftarrow$ iv. p $\rightarrow$q $\Leftarrow$ v. (p $\rightarrow$ Oq) $\Leftarrow$ vi. O(p $\rightarrow$ O q)

d. If John has two cars, he'll either buy a third one or a motorcycle [(b.i)]

e. If John has two cars he'll buy a boat or a motorcycle or may be both [(b.ii)]

There is no strongest construal; the enriched readings do not form a linearly ordered set.

(8) Principles and implementations

a. Alternatives are lexically induced. They are projected upwards as in focus semantics

b. Active alternatives must be used up

c. Operators exploit alternatives to factor in their semantic contribution at designated sites

d. The lexicon: each scalar term is associated with a scale in the lexicon and has a (predictable) variant with no active alternatives.

i. $\|many_{[+\sigma],D}\|^{ALT}$ = {$some_D$, a few_D, $many_D$, $every_D$ }

ii. $\|many_{[-\sigma],D}\|^{ALT}$ = { $many_D$}

(9) Consequences

a. If John has two cars, he'll either buy a boat or a motorcycle

b. O (If John has two$_{[+\sigma]}$ cars, he'll either buy a boat or$_{[-\sigma]}$ a motorcycle) [(9b. ii)]

c. O(If John has two$_{[+\sigma]}$ cars,O (he'll either buy a boat or$_{[+\sigma]}$ a motorcycle)) [(9b.vi)]

d. (OIf John has two$_{[+\sigma]}$ cars, O(he'll either buy a boat or$_{[+\sigma]}$ a motorcycle)) [(9b.i)]

e. Strength Principle: The choice between active alternative/non active alternatives is lexical: choose the entry that fits the context best.
When the same set of alternatives are active, go for the strongest.

(10) Interim conclusions.

An approach based on the hypotheses in (10) together with the strength principle, seems capable of accounting for the Asymmetry Problem.
But we are only considering construnctions in which we deal with one alternative set at a time.

(11) Negative polarity [Kadmon and Landman, Krifka, Lahiri]

a. * There are any students around

b. There are any student around

c. $\exists x \in D_W \, \exists w' \, [student_{W'}(x) \wedge around_W(x)]$

d. Domain widening: (i) "modalize" the N
(ii) activate D-alternatives

e. D = { a, b, c} widest domain

D1 = {a,b} D2 = {b,c} D3 = {a,c}

f. Alternatives: $\exists x \in D_{i,W} \, \exists w' \, [student_{W'}(x) \wedge around_W(x)]$ where $1 \leq i \leq 3$

g. E_C(there are any students around) = $\exists x \in D_{i,W} \, \exists w' \, [student_{W'}(x) \wedge around_W(x)]$
$\wedge \, \forall q \in C \;\; \exists x \in D_W \, \exists w' \, [student_{W'}(x) \wedge around_W(x)] \;<_C q]$

h. $\exists x \in D_W \exists w' \, [student_{W'}(x) \wedge around_W(x)] <_C \; \exists x \in D_{i,W} \exists w' \, [student_{W'}(x) \wedge around_W(x)]$
$= \varnothing$, for if A is entailed by B, A cannot less likely than B

i. $\neg\exists x \in D_W \exists w' [student_{W'}(x) \wedge around_W(x)] <_c \neg\exists x \in D_{i,W} \exists w' [student_{W'}(x) \wedge around_W(x)]$

j. $\|any_{[+\sigma],D}\|^{ALT}$ = {some$_D$, some$_{D1}$, some$_{D2}$,.... }

k. Enrichment via O vs. E depends on what is contextually available: choose O if ALT contain a unique scale (uniqueness principle).

(12) A nice consequence of the uniqueness principle: multiple alternatives.

a. Someone$_{[+\sigma]}$ smokes or$_{[+\sigma]}$ drinks

b. Someone, though not everyone, smokes or drinks, but not both

c. Someone $_{[+\sigma]i}$ [t_i smoke or $_{[+\sigma]}$ t_i drinks]

d. Someone$_i$ [t_i smoke or t_i drinks]

Someone$_i$ [t_i smoke and t_i drinks] everyone$_i$ [t_i smoke or t_i drinks]

Everyone$_i$ [t_i smoke and t_i drinks]

If we wait till the root level, we won't find a unique scale. The only way to obtain a well formed interpretation is to enrich cyclically (i.e. recursively).

e. O [some (one) λx_i O [smoke(x_i) $\vee$ drink(x_i)]

Is O part of an interpretive procedure or is it projected at LF?

(13) Projection of O at LF (Fox 2003)

a. Insert σ

i. σ must have $[+\sigma]$ in its immediate scope

ii. $[+\sigma]$ must be in the immediate scope of σ [cf. the wh criterion]

b. $\| \sigma\ XP\| = O(\|XP\|)$

(14) A mildly complex example

a. $\text{Few}_{[+\sigma]}$ people who smoke $\text{and}_{[+\sigma]}$ drink live up to $80_{[+\sigma]}$

b. There are people who smoke or drink (but not both) and live up to 80

c. There are people who smoke and drink and live to an age close to 80.

d. Some people who smoke and drink do live up to eighty

(15) The syntactic approach

a. * σ [few $_{[+\sigma]}$ people σ who smoke and $_{[+\sigma]}$ drink]i σ [ti live up to eighty $_{[+\sigma]}$]

b. σ σ σ [few $_{[+\sigma]}$ people who smoke and $_{[+\sigma]}$ drink]i [ti live up to eighty $_{[+\sigma]}$]

But how are we going to interpret the sequence of σ's?

You need to consider the arguments of *few* one at a time

b. σ [few $_{[+\sigma]}$ people who smoke and $_{[+\sigma]}$ drink]

i. ignore the alternatives associated with *few*

ii. Generalize pointwise O to other logical types (Rooth)

c. σ σ [few $_{[+\sigma]}$ people who smoke and $_{[+\sigma]}$ drink]i [ti live up to eighty $_{[+\sigma]}$]

i. keep ignoring the alternatives associated with *few*

ii. deal with the alternatives stemming from the second argument of the DP

Then, finally, deal the alternatives associated with *few*.

The following representation looks more plausible

b. σ [few $_{[+\sigma]}$ people who smoke and $_{[+\sigma]}$ drink]i [ti live up to eighty $_{[+\sigma]}$]

But then we need an interpretive approach.

(16) O as part of the interpretive procedure. (Chierchia 1999,2004)

a. $\| \phi \|_S = O_C [\| \phi \|]$, where ϕ is of type t

b. $\|\alpha\ \beta\|_S = \begin{cases} \|\alpha\|_S (\|\beta\|_S), \text{ if } \alpha \text{ is not DE} \\ O_C \|\alpha\|_S (\|\beta\|), \text{ where C is } \|\alpha\|(\|\beta\|)\text{'s scale in } \|\alpha\|(\|\beta\|^{ALT}), \text{ otherwise} \end{cases}$

c. i. A (B) ii. A(O (B)) iii. O A(B))

d. This approach derives the desired interpretation for (15a).

But cannot get *if many students complained we are better off than if all did*

* $(Op \rightarrow q)$

(18) Merging the syntactic and the interpretive approach.

a. Insert σ as in (13a).

b. Define $\| \|_S$ as in (16).

c. $\| \sigma\ XP\| = \iota p[p \in \|\phi\|_S \wedge \forall q\ [\|\phi\|_S(q) \rightarrow p \subset q] \wedge p \neq \varnothing]$

(19) Consequences

The σ–operator becomes a way to enforce the requirement that active alternatives be used up. It also provides a way to obtain disfavored readings.

a.	σ [if John has $\text{two}_{[+\sigma]}$ cars, he is OK]	strong implicature
b.	[σ[if John has $\text{two}_{[+\sigma]}$ cars], he is OK]	weak implicature
c.	If John has $\text{two}_{[-\sigma]}$ cars, he is OK	no implicature
d.	σ [there are any $_{[+\sigma]}$ students around]	contradictory meaning
e.	σ [there aren't any $_{[+\sigma]}$ students around]	sensible meaning

(21) Interaction between SIs and NPIs.

a. everyone who knows any mother with two children should tell them they are eligible for this benefit.

c. **every** (one who knows **any** mother with **two** children)(STE)

d. DE (……NPI……SCAL….)

DE = Downward entailing functor (also a scalar item)

e. ENRICH (every (one who knows any mother with two children))

Where ENRICH = O E or ENRICH = E O

Keep track of alternatives in a way that preserves their order of activation

f. { mother with one child , mother with two children, mother with three children,…}

g. {any_D, any_{D1}, any_{D2},…}

h. { { F1, F2, F3 a1 } { F1, F2, F3 a2 }… }

i. { { any_D mother with one child, any_{D1} mother with one child, any_{D2} mother with one child ….. } { any_D mother with two children any_{D1} mother with two children any_{D2} mother with two children ….. } }

The function determines the structure of the output. Since the function is constituted by D-variants, we get a set of D-variants. So we must apply E first:

f. E every (one who knows any mother with two children) =

every (one who knows any mother with two children)

≤

every(one who knows any mother with two children)

g. Alternatives are obtained by applying E pointwise to the assertion's D-mates

{ E(any_D mother with one child)

E(any_D mother with two children)

E(any_D mother with three children)…….}

h. { any_D mother with one child

any_D mother with two children

any_D mother with three children…….}

At this point we apply O:

i. O(E(every (one who knows any_D mother with two children)) =

j. every (one who knows any_D mother with two children)

$\wedge\neg$ every (one who knows any_D mother with one children)

k. σ [everyone who knows $\text{any}_{[+\sigma]}$ mother with $\text{two}_{[+\sigma]}$ children]$_i$ [t_i should tell them they are eligible for this benefit]

(21) More interaction between SIs and NPIs

a. DE (… scalar term … polar term …)

b. ?? I **doubt** that **every** new doctor has **any** experience.

(22) The D-alternatives induced by *any* come into play before the S-alternatives associated with *every*. So the structure of the alternative set is:

{ {not [every doctor has any_D experience],…., not [some doctor has any_D experience] }

{not [every doctor has any_{D1} experience],…., not [some doctor has any_{D1} experience]}

{not [every doctor has any_{D2} experience],…., not [some doctor has any_{D2} experience]} }

(23) Given the structure of (22) (a set of scales) O must apply (pointwise).

a. O(not [every doctor has any_D experience]) =

not [every doctor has any$_D$ experience] ∧ [some doctor has any$_D$ experience]

b. Alternatives: O(not [every doctor has any$_{D1}$ experience])
O(not [every doctor has any$_{D2}$ experience])
…

(24) a. E(O(not [every doctor has any$_D$ experience])) =

b. **= O(not [every doctor has any$_D$ experience]) $<_C$ O(not [every doctor has any$_{Di}$ experience])** (for any i)

c. = i. not [every doctor has any$_D$ experience] ∧ [some doctor has any$_D$ experience]

⇓ $<_C$ ⇑

ii. not [every doctor has any$_{Di}$ experience] ∧ [some doctor has any$_{Di}$ experience]

(c) fails, for (i) is not stronger than its alternatives

d. Conclusion: σ not [every$_{[+\sigma]}$ new doctor has any$_{[+\sigma]}$ experience] = ∅

e. σ not [every$_{[-\sigma]}$ new doctor has any$_{[+\sigma]}$ experience]

(e) has a sensible interpretation. But it is syntactically ill-formed.

(25) a. I doubt that a new doctor has any experience

b. σ not [a$_{[+\sigma]}$ new doctor has any$_{[+\sigma]}$ experience]

c. O(not [a doctor has any$_D$ experience]) = not [a doctor has any$_D$ experience]

d. Alternatives: not [a doctor has any$_{D1}$ experience]
not [a doctor has any$_{D2}$ experience]
….

f. E(not [a doctor has any$_D$ experience]) = not [a doctor has any$_D$ experience] ∧
not [a doctor has any$_D$ experience] $<_C$ not [a doctor has any$_{Di}$ experience]

(26) a. Few students understood anything
b. Some students understood something

(27) a. E[few(students)](understood anything$_D$)

At he moment in which we apply the DE function few students to its argument understood anything, the alternatives associated with few students will not have come into play yet (cf. again xx)). So we have not choice but to apply E. O comes into play only later.

b. O(E(few(students)(understood anything$_D$))) =
few(students)(understood anything$_D$) ∧
few(students)(understood anything$_D$)] $<_C$ few(students)(understood anything$_{Di}$) ∧
some(students)(understood anything$_D$)]

(28) Conclusions

a. An integrated, uniform theory of NPIs and SIs, governed by the same principles.
   i. Recursive computation of alternatives ($\|\ \|^{ALT}$)and recursive factoring of their contribution ($\|\ \|_S$)
   ii. A syntactically projected abstract operator (σ-trengthen), defined in terms of $\|\ \|_S$.

b. Intervention phenomena have a syntax (agreement) and a semantics (an interpretive clash). The latter makes us see why a certain class of items (strong scalar terms) causes intervention; syntax makes us see why a phenomenon usually optional becomes obligatory

c. Nothing of this appears to be within the range of "globalism".