Do negative polarity items facilitate the processing of decreasing inferences?

Anna Szabolcsi
New York University

joint work with
Lewis Bott, Cardiff University and
Brian McElree, New York University

16th Amsterdam Colloquium, 2007

- Received wisdom: Negative polarity items (NPIs) occur within the immediate scope of operators that support decreasing inferences.
- Expectation: The presence of a licensed NPI facilitates the processing of inferences based on the licensor.
- Conducted a series of experiments to test this.
- How would different NPI theories explain the results?

NPIs are licensed by decreasing operators: Some problems

Weak NPIs are happy with merely Strawson-decreasing licensors (von Fintel 1999)

Only cats ever caught mice.
Only cats caught any mice.

Strong NPIs require anti-additive licensors (Zwarts 1981)

* At most five cats caught mice in weeks.
No cat caught mice in weeks.

Gajewski 2007: In fact, all NPIs are licensed by decreasingness, but…

Weak NPIs only require decreasingness in the truth conditions or assertion.

Strong NPIs require decreasingness that is preserved when all non-truth-conditional or inert coordinates of meaning are taken into consideration.

Standard scalar theories explain only the decreasingness requirement, anyway:

Pick a representative: Kadmon&Landman 1993 – Chierchia 2006 (other versions: Krifka 1996, Lahiri 1998, …)

NPI widens the domain of quantification and requires widening to strengthen the proposition.

This is only possible in a decreasing context.

No cat caught a healthy or ailing mouse (ANY MOUSE) ⇒
No cat caught a healthy mouse (A MOUSE)

Some cat caught a healthy mouse (A MOUSE) ⇒
Some cat caught a healthy or ailing mouse (*ANY MOUSE)

Decreasingness characterizes the inferences that the operator supports.

Decreasingness is a factor in the processing of inferences, independently of the specific meaning of the operator (Geurts & van der Slick 2005).

So, scalar theories raise an **expectation**:
The presence of a licensed NPI highlights the decreasing character of the licensor, and thus facilitates, in one way or another, the processing of decreasing inferences supported by the licensor.

Experiment 1: Inference verification

S1. Our camp is on Staten Island.
S2. Almost <u>every</u>/<u>no</u> camper has
<u>ever</u> / ∅ suffered bruises or caught a cold.
S3. Would it be reasonable to say that almost <u>every</u>/<u>no</u> camper has caught a cold?

S1. Our camp is on Staten Island.
S2. Almost <u>every</u>/<u>no</u> camper has
suffered <u>any</u> / ∅ bruises or caught a cold.
S3. Would it be reasonable to say that almost <u>every</u>/<u>no</u> camper has caught a cold?

Quantifiers contrasted

non-decreasing	decreasing
• Almost every	• Almost no
• Almost everybody	• Almost nobody
• At least five	• At most five
• At least half	• At most half
• More than five	• Less than five
• More than five of	• Less than five of
• Many	• Not many
• Many of	• Not many of
• No less than fifty	• No more than fifty
• No less than five	• No more than five
• Only five	• Very few
• Only five of	• Very few of

Experiment 1: Results

The experiment was sensitive to validity:

Participants discriminated between valid and invalid inferences: they accepted valid inferences much more often than invalid ones.

The experiment had a null result:

Contrary to the expectation, we found no facilitation. The presence of an NPI in S2 did not induce participants to accept more valid inferences or to reject more invalid ones.

Exp 2 & 3: self-paced reading time

S2. Almost <u>every/no</u> camper # has <u>ever</u>/∅ # caught a cold # or suffered bruises. #
S3. Since # almost <u>every/no</u> camper # has <u>ever</u>/∅ # suffered bruises, # the parents are (un)happy, # and ...

S2. Almost <u>every/no camper</u> # has caught a cold # or suffered <u>any</u>/∅ bruises.
S3. Since # almost <u>every/no</u> camper # has suffered bruises, # the parents are …#, and …

Experiments 2 and 3: Results, part1

The experiments were sensitive to NPI licensedness in S2 and to inference validity in S3.

When S2 contained an NPI, participants read the NPI-region and/or the immediately following region significantly slower if the NPI was not licensed.

When neither S2 nor S3 contained an NPI and thus only validity was at stake, participants read the inference region of S3 (*suffered bruises*) and/or the immediately following region significantly slower if the inference was invalid.

Experiments 2 and 3: Results, part 2

No facilitation in S3:

The presence of a licensed NPI in S2 did not speed up the reading of valid inferences in S3. Instead…

Slowdown in S3:

When reading valid inferences, participants significantly slowed down on the inference region if the preceding S2 contained a licensed NPI,

as compared to the case where S2 did not contain an NPI. This effect obtained whether or not the NPI was repeated in S3.

Explain the lack of facilitation and the slowdown

Dissociation + some cost	Association + big cost
The processor does not recognize the NPI licensing property as being the same as the inferential property. So the licensed NPI does not highlight the decreasingness of the operator for the processor. Hence the lack of facilitation.	The processor recognizes the sameness of the NPI licensing property and the inferential property. So there is highlighting, and there is some facilitation for inference accuracy and/or speed.
The processing of the NPI is somewhat costly. Hence the slowdown.	However, the processing of the NPI is very costly. It wipes out any facilitation, and furthermore slows down reading.

Dissociation possibilities

(A) NPI-licensing is syntacticized (say, feature checking), but inference is not.

(B) It is not decreasingness that licenses NPIs, even though there is a big overlap.

(a) non-veridicality licenses many NPIs, ever and any among them (Giannakidou 1998)

(b) +/-interpreted negations are involved (Ladusaw 1992, de Swart & Sag 2002, Postal 2005, Szabolcsi 2004)

+/- interpreted negations

(i) Ladusaw 1992: Reduces n-words in negative concord to NPIs (existentials, not anti-additives). Negation is expressed by a silent operator, whose insertion is licensed by one of the n-words.

Nessuno ha parlato di niente.
no one spoke of nothing
`No one spoke about anything'

(ii) de Swart & Sag 2002: n-words in negative concord are anti-additives forming a polyadic (resumptive) quantifier.

(iii) Postal 2005: NPIs are negation-containing items

Nobody and anybody are alternative surface spell-outs for lexical items with *neg-some* (or *neg-neg-some)*. The choice between no- and any- depends on whether their *neg*'s stay in place or are morphologically deleted by other *neg*'s in the sentence.

One of the arguments for *neg* in any-: secondary triggering.

In years occurs in the context of a local anti-additive. In (3), only any can be that anti-additive.

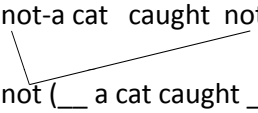
(1) John suspected that no astronauts had gone to Mars in years.

(2) * Nobody suspected that astronauts had gone to Mars in years.

(3) Nobody suspected that any astronauts had gone to Mars in years.

(iv) Szabolcsi 2004 recasts Postal using de Swart and Sag's resumptive quantifier

No cat caught any mouse
not-a cat caught not-a mouse



not (__ a cat caught __ a mouse)
$\neg\exists x,y[cat(x) \& mouse(y)]$ [x caught y]

The flipside of Ladusaw's proposal: reduce NPIs to negative concord.

All decreasing OPs are decomposable into negation + increasing OP. Still, NPI licensing will not highlight the decreasingness of OP for the processor.

NPIs are costly to process -- why?

(A) May be due to something shared by all NPIs.

(B) May be due to something only shared by certain NPIs.

If (A), this will not help with choosing between theories.

Most helpful would be a property that different theories attribute to different NPIs, cf. (B).

Having an even-flavor, thus a scalar semantics, is such a property.

Shared by all NPIs: Lahiri 1998, Chierchia 2006

Only by certain NPIs: Krifka 1995, Giannakidou 2007 (Szabolcsi 2004)

Chierchia 2006

Attributes an even-like flavor to the base meaning of NPIs. This activates a set of domain alternatives and carries the implicature that even the broadest choice of the domain of quantification will make the sentence with the NPI true.

Departing from Grice, implicatures are added and strengthened meanings are calculated recursively, at every step of the sentences's composition.

Domain widening and implicature calculation are plausibly costly.

But NPIs don't always widen the domain…

- N-words interpreted as NPIs (Chierchia 2006)
- Unstressed any applied to unambiguously defined domains (Krifka 1995)

  The empty set doesn't have any proper subsets
- Anymore, either, need, etc. (van der Wal 1999)

  *They *(don't) live here anymore, They *(didn't) laugh either, They need *(not) come early*
- [$_S$ … anti-additive > PPI …] (Szabolcsi 2004)

  I regret / admit that I didn't call someone

  `I regret / #admit that I didn't call anyone'

Chierchia 2006 on scalar items without scalar implicatures

(i) some, many: two versions in the lexicon:
[+σ] strong (with active scalar alternatives) and
[-σ] weak (without scalar alternatives).
Latter replaces the default calculation plus occasional cancellation of the `but not all' implicature.

(ii) NPIs: always [+σ] strong – grammaticized to activate domain alternatives. But the proposition with the widest possible domain needs only to entail its counterparts with particular domains. "Domain widening, as implemented here, is a *potential* for domain widening" (2006:559, emphasis in original).

Chierchia 2006: 557-8

(51) a. *Lexical entry for* any

i. $\|\text{any}_D\| = \lambda P \lambda Q \lambda w\ [\exists w' \exists x \in D_{w'}(P_{w'}(x)) \wedge Q_w(x)]$

ii. $\text{ALT}(\|\text{any}_D\|) = \{\ \lambda P \lambda Q \lambda w\ [\exists w' \exists x \in D'_{w'}(P_{w'}(x) \wedge Q_w(x))] :$
$D' \subseteq D \wedge D'$ is large $\}$

iii. *Any* has an uninterpretable feature $[+\sigma]$.

b. $\|\phi\|_S = E_C\ (\|\phi\|)$, where $C = \|\phi\|^{ALT}$

(50) a. $E_C\ (p) = p \wedge \forall q \in C\ [p \subseteq_c q]$, where $C = ALT$

(53) a. I didn't see any boy.

b. $\sigma \neg$ [I see any boy]

c. $E_C\ (\neg \exists w' \exists x \in D_{w'}\ [\text{boy}_{w'}(x) \wedge \text{see}_w(I, x)])$

d. $\neg \exists w' \exists x \in D_{w'}\ [\text{boy}_{w'}(x) \wedge \text{see}_w(I, x)] \subseteq_c \neg \exists w' \exists x \in D_{i,w'}\ [\text{boy}_{w'}(x) \wedge \text{see}_w(I, x)]$

Processing predictions

The grammaticized activation of domain alternatives plus recursive computation of implicatures predicts that NPIs incur the same processing cost whether or not they tangibly widen the domain.

This contrasts with the prediction of theories that only calculate implicatures for certain NPIs or when strengthening is relevant.

These predictions seem testable

- *Some of the children are in the classroom*
  with the implicature `but not all'

  is costlier to process than

  without the implicature, i.e. `possibly all'.

  (Noveck & Posada, 2003; Bott & Noveck 2004; Breheny, Katsos, & Williams 2005 – argue against default implicature computation)
- Comparable experiments could now be done on negative polarity items.