NIH Public Access
Author Manuscript
J Semant. Author manuscript; available in PMC 2011 May 9.

Published in final edited form as:
J Semant. 2008 November 1; 25(4): 411–450. doi:10.1093/jos/ffn008.

The Effect of Negative Polarity Items on Inference Verification

ANNA SZABOLCSI,
New York University

LEWIS BOTT, and
Cardiff University

BRIAN McELREE
New York University

Abstract

The scalar approach to negative polarity item (NPI) licensing assumes that NPIs are allowable in contexts in which the introduction of the NPI leads to proposition strengthening (e.g. Kadmon & Landman 1993; Krifka 1995; Lahiri 1997; Chierchia 2006). A straightforward processing prediction from such a theory is that NPIs facilitate inference verification from sets to subsets. Three experiments are reported that test this proposal. In each experiment, participants evaluated whether inferences from sets to subsets were valid. Crucially, we manipulated whether the premises contained an NPI. In Experiment 1, participants completed a metalinguistic reasoning task and Experiments 2 and 3 tested reading times using a self-paced reading task. Contrary to expectations, no facilitation was observed when the NPI was present in the premise compared to when it was absent. In fact, the NPI significantly slowed down reading times in the inference region. Our results therefore favour those scalar theories that predict that the NPI is costly to process (Chierchia 2006), or other, non-scalar theories (Ladusaw 1992; Giannakidou 1998; Szabolcsi 2004; Postal 2005) that likewise predict NPI processing cost but, unlike Chierchia (2006), expect the magnitude of the processing cost to vary with the actual pragmatics of the NPI.

1 INTRODUCTION

Negative polarity items (NPIs) are expressions whose occurrence is restricted to the immediate scope of certain operators, called their licensors. *Ever* is an NPI licensed, among other operators, by *none/few/ at most two of the books*:

1. *All of the books have ever been borrowed.

   *Some of the books have ever been borrowed.

   *Most of the books have ever been borrowed.

2. None of the books have ever been borrowed.

   Few of the books have ever been borrowed.

   At most two of the books have ever been borrowed.

Some other NPIs are *yet, at all, in weeks, sleep a wink, much* and *any*. Different NPIs are licensed by somewhat different operators (Zwarts 1981; Giannakidou 1998; De Decker *et al*. 2005), but the classical view, going back to Ladusaw (1980), is that the licensors of NPIs in

---

© The Author 2008. Published by Oxford University Press. All rights reserved.

ANNA SZABOLCSI, Department of Linguistics, New York University, 726 Broadway, New York NY 10003, USA, anna.szabolcsi@nyu.edu.

English are at least monotone decreasing.[1] A monotone-decreasing (DE) operator reverses the ordering in its domain. If $A \leq B$ and Op is decreasing, then $Op(B) \leq Op(A)$. One special case is where $A \leq B$ represents 'A is a subset of B' and $Op(B) \leq Op(A)$ represents 'Op(B) entails Op(A)'. For example, both inferences below illustrate the fact that 'few of the books' is a decreasing operator:

(3) If $VP1 \subseteq VP2$, DE-Op(VP2) ⇒ DE-Op(VP1)

Given that the things that have been borrowed by children are a subset of those that have been borrowed,

Few of the books have been borrowed ⇒

Few of the books have been borrowed by children

(4) DE-Op(VP1 or VP2) ⇒ DE-Op(VP1); DE-Op(VP2)

Given that things that have been borrowed are a subset of those that have been borrowed or used in the reading room,

Few of the books have been borrowed or used in the reading room ⇒

Few of the books have been borrowed; Few of the books have been used in the reading room

Therefore, (5) is a correct definition and (6) a correct descriptive generalization.

(5) An operator is decreasing iff it supports inferences from sets to subsets.

(6) Correlation between NPI-licensing and inferences:

Negative polarity items occur in the immediate scope of operators that support inferences from sets to subsets.

The central question of this paper pertains to the psychological status of this abstract grammatical generalization. In particular, we ask whether human sentence processing operations (henceforth, *the human processor* or *the processor*) recognize that the distribution of NPIs is governed by the same property that supports inferences from sets to subsets.

1.1 The association model

The correlation between NPI licensing and inference to subsets need not be just a descriptive coincidence; it may be essential and explanatory. For example, the main hypothesis in Dowty (1994) is formulated as follows (his (27)):

> **Hypothesis**. Given that (i) ↑ M and ↓ M inferences are a very significant pattern of natural language reasoning, and (ii) the distribution of NPIs (and NC) is (almost)

---

[1] There are well-known problems with the 'at least monotone decreasing' generalization in English. For example, *only* licences various NPIs, although it is overall non-monotonic:

(i) Only thrillers have ever been borrowed.

Von Fintel (1999) argues that *only thrillers* satisfies a weaker notion that he dubs Strawson decreasingness. This is restricted to contexts where relevant presuppositions are borne out. See Giannakidou (2006) for critical discussion. Gajewski (2008) unifies weak and strong NPIs as both being licensed by decreasingness and differing, in that weak NPIs only care about the truth conditions/ assertion, while strong NPIs require decreasingness that is preserved when all non-truthconditional/ inert coordinates of meaning are taken into consideration. On the other hand, Postal (2005) observes that *zero books* is decreasing but does not license NPIs:

(ii) *Zero books have ever been borrowed.

Finally, there are subtler problems, discussed in Zwarts (1995) for English and Giannakidou (1998) for Modern Greek. This paper will gloss over these matters and investigate a plain version of the classical view.

> coextensive with logically ↓ M contexts, we can hypothesize that one important reason for the existence of NPI and NC marking is to directly mark positions syntactically which are subject to ↓ M inferences. (↓ *M and* ↓ *M stand for 'monotonically increasing' and 'decreasing', respectively, and NC stands for 'negative concord'—AS, LB and BM*.)

The assumption of an essential correlation also applies to scalar accounts of NPI licensing (Kadmon & Landman 1993; Krifka 1995; Lahiri 1997; Chierchia 2006). Glossing over the differences between these accounts, let us say that the NPI widens the domain of quantification and, moreover, comes with a requirement that domain widening should strengthen the claim as compared to what the use of a plain indefinite would convey. For example, the plain indefinite *(some) readers* is taken to quantify over typical readers, and the NPI *any readers* to quantify over a widened domain of either typical readers or marginal readers, such as paid reviewers or people who just browse the book. Now consider (7) and (8).

(7) **a.** Many of the books had (some) typical readers ⇒
Many of the books had (some) typical or marginal readers
b. *_Many_ of the books had _any_ readers.

(8) **a.** Few of the books had (some) typical or marginal readers ⇒
Few of the books had (some) typical readers
b. $^{\text{OK}}$_Few_ of the books had _any_ readers.

A proposition *p* is stronger than *q* if *p* entails *q*. In an increasing context such as (7), having typical readers entails having either typical or marginal readers, so by widening the domain we weaken the claim instead of strengthening it. *Some* cannot be replaced by *any*. In contrast, (8) is a decreasing context and the entailment relations are reversed. Few of the books having typical or marginal readers entails, and is thus stronger than, few of them having typical readers. Here, *some* can be replaced by *any*.

If this account is correct, then it is quite plausible for the processor to recognize that NPI licensing and supporting inferences to subsets are the same property. Let us call this the association model.

The association model presupposes that decreasingness is a factor in the processing of inferences, as do theories within formal semantics (Saánchez-Valencia 1991; van Benthem 1991; Bernardi 2002; Fyodorov *et al*. 2003; Altman *et al*. 2005). The only psychological data that have tested this claim, however, is Geurts (2003a) and Geurts & van der Slik (2005). Geurts argued that if montonicity plays a role in human reasoning, then syllogisms that are complex according to a monotonicity-based logic should show low rates of success in human reasoning tasks. He therefore calculated the number of steps required to solve standard syllogisms according to a set of monotonicity-based inference rules, and compared the complexity of the problem with human performance. He found that the monotonicity-based complexity did indeed correlate with performance (but see Newstead 2003, and Geurts 2003b, for commentary). Geurts and van der Slik came to a similar conclusion by demonstrating that participants find sentences with two quantifiers of the same monotonicity profile easier to interpret than sentences with mixed monotonicity profiles. Thus, what evidence there is on processing and monotonicity agrees with the association model.

Given the existing evidence on NPI licensing, it seems almost like a foregone conclusion that the processor should recognize the relationship between monotonicity and NPI licensing. Indeed, when we started this project, our intention was to investigate how this recognition occurred, rather than whether it occurred. As is often the case, however, expectations did not match results and we were unable to find evidence for the recognition. We therefore describe other linguistic models, which do not predict such a relationship.

1.2 The dissociation model

The essential relation (e.g. scalar) account of NPI licensing is also compatible with the processor failing to recognize the sameness of properties. It is possible, for example, that the NPI licensing effect of decreasing operators is compiled into the syntax, whereas inferencing is performed purely model theoretically. It could be argued that NPI licensing is purely syntactic. For example, NPIs have a [−de] feature that makes the sentence ungrammatical unless it is deleted in construction with a licensor, which bears a [+de] feature. On this view, the set of operators that bear the syntactic [+de] feature may coincide with the set of semantically decreasing ones, but this fact has no significance in the actual licensing process. In contrast, inferences could be computed purely in a model-theoretic semantic manner, not in terms of syntactic features.

The processor may dissociate NPI licensing from inferences for another reason. It may be that the correlation between licensing and inference to subsets holds true, but it is not essential for how NPI licensing works. For example, Zwarts (1995) and Giannakidou (1998) propose that non-veridicality, as opposed to decreasingness, is the semantic property that licenses certain polarity-sensitive items. The baseline definition is that an operator Op is veridical if Op(p) entails p, and non-veridical if Op(p) does not entail p. Giannakidou introduces further qualifications to narrow down the set of licensors; see Simons (1999) for a review. Relevant to us is that all decreasing operators are non-veridical (e.g. *never*), although not all non-veridical operators are decreasing (e.g. *or*), as illustrated in (9) and (10):

(9) Mary never ate pizza ⇒ Mary never ate pizza with anchovies (*never* is decreasing)

Mary never ate pizza ⇏ Mary ate pizza (*never* is non-veridical)

(10) Mary is sad or tired ⇏ Mary is sad; Mary is tired (*or* is non-veridical)

Mary is sad or tired ⇏ Mary is very sad or very tired (*or* is not decreasing)

Suppose now that some NPI is licensed by non-veridical operators. The presence of such a licensed NPI would not highlight the decreasingness of the licensor even if the licensor happened to be decreasing, because its decreasingness would not be essential to licensing.

Finally, there is a family of theories that identify interpreted or uninterpreted negatives as the crucial factors in NPI licensing. Ladusaw (1992) assimilates Romance negative concord to NPI licensing, arguing that n-words (*nessuno*, *nadie*, *personne*, etc.), as well as verbal negation in Romance languages, are NPIs, and their licensor is an overt or silent anti-additive item.[2] de Swart & Sag (2002) recast this analysis with n-words interpreted as anti-additive quantifiers that are absorbed into a single polyadic quantifier.

(11) Nadie no he visto nada.

no one not saw nothing

‘No one saw anything’

---

[2] Op is anti-additive if it bears out this de Morgan law: Op(a or b) iff Op(a) and Op(b). *No one, never, without* are anti-additive. If Op is merely decreasing, the biconditional holds left to right but not right to left.

Postal (2005) and Szabolcsi (2004) propose the flip-side account and assimilate NPI licensing to negative concord. More precisely, according to Postal, NPIs are not lexical items in need of licensing. Instead, surface forms like *no one* and *anyone* are alternative morphologies that spell out the combination of an underlying indefinite and one or more negations,[3] the choice depending on whether the negations are left alone or cancelled by other negations in the sentence. Szabolcsi recasts Postal's proposal along the lines of de Swart and Sag: both the NPI and the licensor have a negation component in their lexical semantics; these negations are factored out to form a polyadic negative quantifier (see Szabolcsi 2004: 433–6 and 447–50). The following example involves some simplifications that are irrelevant to present purposes:

(12) not-many books had not-some readers = underlying

not (_ many books had _ some readers) = negations factored out

Few books had any readers. = spell-out

As all decreasing operators either are negations or can be decomposed into a negation plus an increasing operator within its scope (for *little* as degree negation, see Heim 2006), this analysis is fully compatible with the correlation between NPI licensing and decreasingness. But the processor will have no reason to recognize that the NPI licensing property of *few books* is coextensive with the one that supports inference to subsets. In other words, dissociation, while not strictly necessary, is quite plausible.

1.3 Overview of the experiments and empirical predictions

The association and dissociation models make different predictions regarding the effects of an NPI on inference processing. We conducted three experiments to discriminate between the models. In all three of the experiments, participants verified whether sentences involving quantifiers were entailed by other sentences. For example, the participant might read 'None of the colleagues have *ever* sent flowers or cards', and judge whether 'None of the colleagues have sent cards' was true. The crucial manipulation was whether the disjunction sentence contained an NPI. The association model, combined with Geurts' findings, makes a strong prediction that the occurrence of an NPI within the licensing domain of a decreasing operator should facilitate processing of inferences from sets to subsets. Below we describe how we arrive at this prediction.

According to the association model, facilitation should arise for two reasons. First, the presence of the NPI is assumed to provide additional information as to the decreasing monotonicity of the context. As Geurts and van der Slik write when reviewing the literature on NPIs, 'In effect, a NPI serves to *signal* that the environment in which it occurs is [decreasing] … ' (p. 240, emphasis in the original, in Geurts 2003a: 101; Geurts & van der Slik, 2005). With more evidence contributing to the monotonicity classification decision, the classification process will be less susceptible to processing noise, which may mask relevant contextual information. We would expect this to translate into either more accurate monotonicity judgments or quicker processing times, depending on how participants elect to balance speed and accuracy constraints in their classification judgments. If the monotonically decreasing character of the context is more easily determined, and this information is used in the inference verification process, inference verification will show facilitation relative to a situation in which monotonicity decisions are difficult to determine.

---

[3] One of Postal's strong descriptive arguments for the claim that *any* forms contain a lexical negation comes from the phenomenon known as 'secondary triggering'. The NPI *in years* requires a clausemate anti-additive licensor. (i) is bad because it does not have one.

i. *Nobody suspected that astronauts had gone to Mars in years.

ii. Nobody suspected that no astronauts had gone to Mars in years.

iii. Nobody suspected that any astronauts had gone to Mars in years.

In (ii), the clausemate anti-additive licensor is *no astronauts*. In (iii), only *any astronauts* can be the licensor, but then it must contain an anti-additive operator, just like *no astronauts* does. Den Dikken (2006) generalizes Postal's account to the Dutch NPI *hele* 'whole' and recasts the analysis in syntactic terms similar to Ladusaw's, completing the circle.

Second, facilitation could arise if the NPI forces computation of the monotonicity context prior to the point at which inferences are needed. If there is no NPI in a quantifier sentence, there is no guarantee that the monotonicity of the context will be computed until inference verification is required (indeed, there are good arguments on psychological and linguistic grounds that sentences are minimally processed until discourse constraints force otherwise, see, e.g. McKoon & Ratcliff 1992; Koller & Niehren 1999; Ferreira & Patson 2007). Processing time for the monotonicity calculation will therefore be delayed until inferences need to be verified. This is not the case when an NPI is present in a sentence, however. Since recognizing decreasingness is assumed to be a necessary part of the NPI licensing computation, a sentence with an NPI cannot be processed without determining decreasing monotonicity. Accordingly, when inferences need to be verified later in the discourse, the monotonicity computations will have already been completed, and thus the valid inferences can be determined relatively easily.

The association model predicts that facilitation of inferences arises because the NPI should highlight the monotonicity properties of the quantifier and because it forces the monotonicity computations prior to the inference verification stage of processing. In contrast, if NPI licensing is not essentially linked to the monotonicity computations, as in the dissociation model, no facilitation of inference processing would be expected. These predictions were tested using a reasoning task and two self-paced reading tasks, presented below in section 2. In the case of each experiment, the introductory section explains the hypotheses and how the materials were constructed. This is followed by a technical presentation of the method and the results. The final section for each experiment is the discussion of the extent to which the effects observed in that experiment supported the hypotheses. The more general linguistic interpretation of the results is reserved for the general discussion in section 3.

2 EXPERIMENTS

2.1 Experiment 1

Experiment 1 was a reasoning task in which participants judged whether a decreasing inference was supported by the preceding discourse. Participants read two-line vignettes, followed by a question. The first line set the general context, and the second line involved a verb phrase disjunction within the scope of a subject quantifier. Two types of quantifiers were used, decreasing and non-decreasing. When the quantifier was decreasing, the combination supported an inference to one of the disjuncts within the scope of the same quantifier, but when the quantifier was non-decreasing, the inference was not supported. The question at the end of the vignette explicitly asked about the inference to one of the disjuncts. Crucially, half the items included an NPI (*ever* or *any*), while the other half did not. If the association model were correct, we would expect the inferences that included the NPI to be verified more accurately than those that did not include the NPI. We now elaborate on the structure of the materials used in the task and the experimental design.

Participants answered 48 reasoning problems constructed from different vignettes. Each vignette existed in four conditions: (a) decreasing quantifier with an NPI; (b) decreasing quantifier without an NPI; (c) non-decreasing quantifier with an NPI; (d) non-decreasing quantifier without an NPI. Inferences from DE quantifiers (conditions a and b) supported inferences to subsets (i.e. the inference presented in S2 was *valid*), whereas inferences from non-decreasing quantifiers (conditions c and d) did not (inferences were *invalid*). Conditions (a) and (c) contained NPIs, but not (b) and (d) (note that the NPI in (c) is unlicensed for most non-decreasing quantifiers). For each vignette, pairs of comparable quantifiers were used for the DE and the nondecreasing conditions, respectively. After participants had read the context sentence and the disjunction sentence, they answered an inference question in the format, ‘Would it be reasonable to say that Quantifier(VP2)?’ Table 1 shows two examples of the stimuli. The context sentence, displayed in all conditions, is shown in the top row. Sentences displayed in the different conditions are shown in the following rows, and the question appears in the final row. Note that the inference invariably used the second disjunct so as to avoid repetition of a contiguous stretch of the previous sentence. The correct answer was ‘yes’ for conditions (a) and (b) and ‘no’ for (c) and (d).

There were 48 vignettes, constructed so as to support *ever* as the NPI, or to support *any* (referred to in the discussion below as *ever* items and *any* items, respectively, regardless whether they are used in the NPIpresent or the NPI-absent condition). The complete list of the quantifier pairs is shown in Table 2. Each participant saw only one occurrence of each quantifier and only one use of each particular vignette.[4]

If the association model is correct and there is an essential link between decreasingness and NPIs, participants should find version (a), in which an NPI is present in the disjunction sentence, easier to process than version (b), in which no NPI is present. As described in section 1.3, the NPI should highlight the monotonicity context and make it easier for participants to verify the inference. If there is not an essential link between NPI licensing and decreasingness, there should be no difference between the versions. Conditions in which there is no NPI present can be used to measure baseline performance. Participants should respond positively to valid inferences without the NPI (b) and negatively to invalid inferences without the NPI (d). We included condition (c)—with an unlicensed NPI—because we wished to establish whether participants were processing the NPI independently from whether the NPI affected processing of the inference. Under conditions in which participants showed no effects of reading an unlicensed NPI, an association model would not necessarily predict that an NPI would facilitate processing. Condition (c) allows us to eliminate the possibility that participants were merely ignoring the NPI during the task. (This possibility is more likely in later experiments in which participants merely read the sentences without having to explicitly judge the veracity of inferences.)

2.1.1 Method

Participants: Twenty-eight New York University students participated for course credit or payment. Participants were randomly allocated to one of four experimental groups (see below).

---

[4] There were several reasons for these quantifier choices. Speakers generally find modified quantifiers more natural and plausible, which is important in preventing the task from having an IQ-test feel. Regarding specific quantifiers, *almost no(body)* is strictly speaking non-monotonic. However, people readily accept inferences like this:

i. Almost no campers have had a sunburn or caught a cold ⇒

Almost no campers have caught a cold.

We believe the reason is this. Indeed, it may be that some campers had sunburns but no campers caught colds, in which case *almost* is not justified in the conclusion. However, given that campers in a camp usually number in the hundreds and sunburns and colds are equally probable and unremarkable, *almost no* essentially means that the number of incidents was negligible. People freely assume that both kinds of events occurred and if for some reason one did not, it is not important. This is different from inferences to arbitrary subsets, as in Almost no campers had a sunburn or sprouted a second head ⇏ Almost no campers sprouted a second head.
The quantifiers involving *only* license the NPI although they are used in condition (c) (they are non-monotonic). These items were therefore removed for the analysis of the effect of the unlicensed NPI and reinstated for all other analyses.

Design and stimuli: Forty-eight items were constructed, 24 capable of supporting *ever* and 24 capable of supporting *any*. Each item was three lines long and included a context sentence (sentence 1, S1), a *disjunction* sentence that used *or* (S2), and an *inference* question (S3). The disjunction sentence always began with a quantifier, and included an NPI in the (a) and (c) conditions. For the *ever* items, the NPI occurred prior to the disjunction and scoped over both disjuncts. For the *any* items, the NPI occurred in the first disjunct and therefore only scoped over the first disjunct. The inference question always began with the phrase, 'Would it be reasonable to say that', followed by the quantifier, followed by the inference. The NPI was not included in the inference sentence because we wished the inference sentence to be identical across all conditions. A selection of the items is shown in the Appendix, where the first two sentences in each item were used as the first two lines in this experiment, and the third sentence was transformed into the question (although note that the *any* items presented in the Appendix correspond to the self-paced versions used in Experiment 3, not the versions used in Experiment 1).

Four experimental groups of items were created, so that each participant would see items from all four conditions, but see no single item more than once. Items were counterbalanced across groups, so that all items occurred equally often in all four conditions across the experiment.

Participants also saw 24 filler items based around simple deductive inferences, 12 of which were true and 12 of which were false. Filler items were of the same general form as the experimental items but did not vary over experimental groups. For example, a true filler item was, 'John looked in the fridge./He found ham and cheese./Would it be reasonable to say that John found food in the fridge?'

Procedure: Participants read the sentences from a computer screen. A fixation cross first appeared, followed by the first sentence of the item. Participants pressed the enter key to advance to the second and third sentences, respectively. After reading the third sentence, they pressed a key corresponding to 'yes' or 'no' to indicate their response to the question.

2.1.2 Results—Accuracy on the filler items was at ceiling levels. For the valid inference filler items, average proportion 'valid' responses was 0.96, whereas for the invalid items, average proportion 'valid' responses was 0.05 (hence, proportion correct = 0.95). Thus, participants understood the instructions and experienced no difficulty in answering simple deduction questions.

Average proportion 'valid' judgments for the four experimental conditions are shown in Table 3. Valid items (a and b) were generally answered correctly, but participants also believed that a large proportion of the invalid items (c and d) were valid inferences (hence, proportion correct scores were low for the invalid items: M = 0.51 and M = 0.58, respectively). Nonetheless, there was a robust difference between valid and invalid items whether the NPI was present or not, (a) v. (c), $F(1,24) = 39.27$, $P < 0.0005$, (b) v. (d), $F(1,24) = 40.2$, $P < 0.0005$ (we used counterbalancing group as a blocking factor, as we did in all the analysis of variances (ANOVAs) reported in this article). Participants therefore processed the quantifiers sufficiently deeply to discriminate between valid and invalid inferences, but they experienced some difficulty in rejecting invalid items.

We next consider the effect of the NPI. We hypothesized that the presence of the NPI would facilitate participants in judging whether the inference was valid. For the valid items, the addition of the NPI should therefore lead to a greater proportion of valid responses, i.e. (a) > (b). Table 3 shows that this was not the case. For the valid items, the means are essentially equal, $M_{a–b\ diff} = −0.0012$, SD = 0.13, 95% confidence interval (CI) participants (one tailed): $-0.042 \leq M_{diff} \leq 0.040$, 95% CI items (one tailed): $-0.036 \leq M_{diff} \leq 0.033$. We can therefore be confident that the NPI improved performance by a maximum of 4% for the valid items.

Accuracy on the invalid items appears to have been worsened by the introduction of the NPI. To examine this in more detail, we broke the invalid items into the *ever* set and the *any* set and performed a repeatedmeasures ANOVA with presence of NPI and type of NPI as factors. This revealed marginal effects of the presence of the NPI, $F_1(1,27) = 3.017$, $P = 0.094$, $F_2(1,46) = 4.21$, $P = 0.043$ and of the interaction, $F_1(1,27) = 5.05$, $P = 0.033$, $F_2(1,46) = 2.94$, $P = 0.093$, such that the addition of *any* worsened accuracy more than the addition of *ever*. Unfortunately, the drop in accuracy in condition (c) coincided with a progression towards chance. Hence, it is difficult to know whether participants believed more inferences were valid, or whether more participants responded at chance because they found the sentence with the unlicensed NPI too difficult to interpret.

Correct judgments of the valid items were quite high ($M = 0.87$) and we were concerned that ceiling effects might be obscuring a facilitatory effect of the NPI. We therefore examined the individual participant data to check that incorrect scores were not restricted to a minority of participants. However, only 5 out of 28 participants had perfect scores and there was room for improvement in the remaining 23 participants. The non-ceiling participants scored $M = 0.85$ v. $M = 0.84$ in the valid inference, NPI-present and NPI-absent conditions, respectively, which is not a reliable difference, $t < 1$. Moreover, performance on items in which participants had to reject the inference as being invalid was low ($M = 0.58$ correct), and the high discrepancy between these scores indicates that high performance on the valid inferences may have been due to a confirmation response bias. In short, our concerns about ceiling effects were unfounded.

We also analysed response times to establish whether participants responded more quickly to valid inferences with the NPI (condition (a)) than valid inferences without the NPI (condition (b)). Mean response times to correctly answered inferences were almost identical, however, $M_a = 2.69$ s ($SD = 0.86$) v. $M_b = 2.66$ s ($SD = 0.97$), and no significant differences were apparent, 95% CI (participants, two tailed): $-0.19 \text{ s} \leq M_{a-b} \leq 0.24$ s, 95% CI (items, two tailed): $-0.26 \text{ s} \leq M_{a-b} \leq 0.27$. Thus, the RT analysis was consistent with the response proportion analysis. We did not analyse the invalid inferences because there was an insufficient number of correct responses.

Finally, we analysed whole sentence reading times for the sentence that contained the NPI (S2). In particular, we wished to establish whether participants had processed the NPI at all, given that we observed no facilitatory effect in either response proportions or response times. We compared condition (a) with condition (c), both of which contained an NPI but in (a) the NPI was licensed, whereas in (c) it was not. Participants read the unlicensed condition much more slowly than the licensed condition, $M = 4.67$ s ($SD = 1.45$) v. $M = 4.05$ s ($SD = 1.07$), $F_1(1,27) = 18.11$, $P < 0.005$, $F_2(1,46) = 10.81$, $P < 0.005$, but there was no main effect of NPI type (*ever* v. *any*), $F_s < 1$, or interaction, $F_s < 1$. Thus, we can be sure that participants processed the NPI and the quantifier in sufficient depth that they were sensitive to whether the NPI was licensed.

2.1.3 Discussion—We tested participants on inferences that involved either decreasing or non-decreasing quantifiers. We hypothesized that when an NPI was included in the text, participants should have been more accurate in judging that inferences involving decreasing quantifiers were valid, compared to contexts in which the NPI was absent.

Participants generally accepted valid inferences, but experienced difficulty in rejecting invalid inferences. Nonetheless, there was a robust difference between validity judgments in the two conditions. The presence of the NPI, however, failed to significantly facilitate correct acceptance of the valid inference. Indeed, accuracy on the valid inferences was identical in the NPI-present and NPI-absent conditions, and we can be 95% confident that the NPI facilitated processing by at most 4%. One potential reason for the absence of an effect of the NPI is that participants might not have processed the NPI to a sufficient depth when they read the sentences. However, this is an unlikely explanation because we observed marginally lower accuracy when the NPI was present in the non-decreasing quantifier condition (c) (where it was generally unlicensed) than when it was not present (d), and participants spent an extra half a second longer reading sentences from condition (c) than from condition (a). Hence, there is no doubt that participants were paying attention to the quantifier–NPI combination, yet they appeared not to use this information when evaluating the inferences.

This experiment demonstrated that the presence of an NPI does not substantially facilitate the accuracy with which people can make inference judgments. Nonetheless, facilitation might reduce processing time without improving accuracy. Although we did not observe response time differences on the inference sentences, we could only observe reading times for the whole sentence, which included the ‘Would it be reasonable’ section and other parts of the sentence that might obscure a small processing time facilitation. We therefore conducted reading time experiments in which we could measure reading times on specific regions of the sentence.

2.2 Experiment 2

Experiment 2 tested whether the addition of the NPI reduced reading time when participants read a version of the inferences presented in Experiment 1. The inferences were embedded in short vignettes containing four sentences: a context sentence, a sentence containing a quantifier and the NPI (in the appropriate condition), a sentence containing the inference and a sentence closing the vignette. The sentences were divided into regions and participants advanced from region to region by pressing a key. Table 4 shows one of the items, the ‘Staten Island’ vignette. The slashes indicate self-paced reading regions.

The first two sentences of each item were the same as those used in Experiment 1. The third sentence always began ‘*Since* Quantifier(VP2)’ and contained the inference to the second disjunct that was expressed as the question in Experiment 1. (Notice that S3 did not repeat a contiguous stretch of S2.) We assumed that the most felicitous use of *since* (‘seeing that’) is one where the content of its complement is given information, for example, inferable from what has just been conveyed (see also Verbrugge & Schaeken 2006). Using *since* should therefore force participants to verify the inference, obviating the need for an explicit question.[5]

We predicted that if the presence of the NPI facilitated processing, reading time on the inference region of the valid inference (‘has/have caught a cold’) would be quicker when the NPI was present than when it was absent ((a) v. (b)). We predicted effects on the inference region because this is the earliest point at which the NPI could facilitate inference generation. Since the language processor is known to act incrementally, the inference region seemed a more plausible site for facilitation than later regions. Note that S3 was identical across conditions (a) and (b) since the NPI only occurred in S2, the premise with the disjunction. Participants should also be slower in the inference region when reading invalid inferences with no NPI, compared to valid inferences with no NPI (S3.3, (d) v. (b)), and they should be slower reading an unlicensed NPI than a licensed NPI (S2.2, (a) v. (c)). 'S3.3' denotes region 3 of sentence 3; 'S2.2' region 2 of sentence 2, etc.

---

[5] People seem to find it easier to make both scope judgments and inferences when anaphora and non-metalinguistic questions are used, rather than when metalinguistic reasoning is tested (see Paterson *et al*. 1998; Tunstall 1998; Szabolcsi 2006).

2.2.1 Method

Participants: Thirty-two New York University students participated for course credit or payment. None had completed Experiment 1. Participants were randomly assigned to four experimental lists.

Stimuli and design: Items were the same as the *ever* items generated for Experiment 1,[6] except for additional material presented at the end of S3 and S4, and the use of *since* instead of a question. To ensure that all the sentences were felicitous, the additional material in S3 varied slightly with the quantifier (e.g. 'the parents are happy,/ and they praise the counselors/' *vs*. 'the parents are unhappy,/ and they blame the counselors/'). These differences occurred after the regions of interest and so they cannot affect the conclusions of the experiment. We also added a question at the end of each item to maintain the participant's attention. Half of these questions were designed to have true answers and half were designed to have false answers. The items were divided into the same four lists as in Experiment 1. Items were divided into regions through which participants advanced at their own pace. The first sentence was divided into three regions of approximately the same length. For the second sentence, the division of the sentence varied depending on the structure of the item. For all items, the first region contained the quantifier phrase, and the second region contained the NPI (in the (a) and (c) conditions). This region was designed so that it contained an equal number of words across all the items (two words plus the NPI), so that variation across items was minimized on this region. The remaining sentence was divided into two or three regions depending on the length.

The third sentence always included five regions. The first was *since*; the second was the quantifier phrase; the third was the second disjunct of S2 followed by a comma (the inference region) and the fourth and fifth were concluding phrases.

We also included 40 filler items that were of similar length to the experimental items but did not include NPIs or inferences based on the quantifier.

Procedure: At the start of each trial, each participant saw the first sentence of an item appear as a row of underscores, each underscore representing a missing letter. The first set of words then appeared in the first region, and the participant advanced onto the next region by pressing space. The words of the current region disappeared and the next set arrived in place of the underscores. After the participant had read the first sentence, the next sentence appeared below it and the old sentence disappeared. On completing all four sentences, the comprehension question appeared and the participant pressed a key corresponding to the yes/no response. After the response, the first sentence of the next trial appeared and the participant proceeded as before.

---

[6] We also included a set of items with *any* as an NPI. However, when we analysed these items in isolation, we found that participants were only marginally sensitive to the validity of the inference and to the presence of the unlicensed NPI, and there were no experimentally interesting significant differences. Since very little can be concluded from these items, and to aid the exposition, we report only the analysis of the *ever* items (although the experiment was analysed on the basis that both sets of items were tested to maintain appropriate family-wise error rates, and the reporting of *P* values and degrees of freedom reflect this). The complete analysis is available on request.
One reason why the *any* items might have behaved differently to the *ever* items is that *any* did not modify the proposition on which the inference is based (*any* modified the first disjunct but we tested on the second disjunct), unlike *ever* in the *ever* items. However, this could not explain why we did not get strong effects of the unlicensed NPI or the inference, so we prefer to remain agnostic about possible differences between the NPIs and about the role of the scope position of the NPI.

2.2.2 Results

Data preprocessing: We removed all RTs that were more than 4 SDs away from the mean of their region. We also removed all data associated with items that were incorrectly answered because we had no guarantee that participants had read the items in these cases. For the current experiment, 7% of the data were removed because of incorrectly answered comprehension questions. A further 4% were removed as outliers from S2, and 4% from S3.

Sentence 2: S2 contained the NPI in conditions (a) (licensed) and (c) (generally unlicensed). Since we observed only a marginal effect of the NPI on accuracy in Experiment 1, we analysed S2 to establish that participants were processing the NPI in this experiment. Since condition (c) generally contained an unlicensed NPI, participants should spend longer reading these sentences than condition (a).

Figure 1 displays the within-subject difference in reading time between conditions (a) and (c) as a function of the sentence regions. We consider regions 2, 3 and 4 because prior to region 2, the NPI had not been presented. Participants were quicker reading the unlicensed NPI than the licensed NPI but slowed down in the following region. This is confirmed by an interaction between region and NPI license condition ((a) v. (c)), $F_1(2,62) = 6.12$, $P = 0.004$, $F_2(2,30) = 6.03$, $P = 0.006$.

Sentence 3: S3 contained the inference. In conditions (a) and (b), the inference was valid, whereas in (c) and (d) it was invalid. We expected participants to have longer RTs when reading invalid inferences than valid inferences independently of the effect of the NPI. We therefore compared conditions (d) and (b) on the inference region (S3.3). The upper line of Figure 2 shows the within-subject difference between conditions (d) and (b) of S3. On the inference region, the invalid inferences were read significantly more slowly than the valid inferences, $t_1(31) = 3.79$, $P = 0.001$, $t_2(23) = 0.004$. Furthermore, the effect was not present on regions prior to the inference region, $t_1(31)$s < 1.66, Ps > 0.10, $t_2(23)$s < 1.77, Ps > 0.09. This was confirmed by an interaction between region and inference type, $F_1(2,56) = 5.22$, $P = 0.008$, $F_2(2,46) = 4.92$, $P = 0.012$, when regions S3.1, S3.2, and S3.3 were included in a repeated-measures ANOVA. Thus, the increased reading times on the inference region must be due to participants evaluating an invalid inference, and not merely due to participants reading different quantifiers across the two conditions.

We hypothesized that if the NPI facilitated inference verification, RTs for valid inferences should be lower when the preceding sentence had a licensed NPI. We therefore compared valid inference reading times when the preceding sentence contained the NPI to when it did not, i.e. (a) v. (b). The lower line of Figure 2 shows the within-subject difference between the two conditions. Interestingly, and counter to an NPI-facilitating hypothesis, conditions involving an NPI were more difficult to read than those without the NPI. This is illustrated by the predominately negative values in the figures. There was a robust effect on the inference region, $t_1(31) = 4.74$, $P < 0.0005$, $t_2(23) = 2.95$, $P = 0.007$. Furthermore, this effect is greater on the inference region than on the regions prior to it, $F_1(2,56) = 7.60$, $P = 0.001$, $F_2(2,46) = 4.76$, $P = 0.013$.

In summary, there was no evidence that the NPI facilitated processing. Indeed, we observed a robust increase in RTs in the context of the NPI.

2.2.3 Discussion—The primary goal of this experiment was to determine whether the presence of an NPI facilitated inference making in decreasing contexts. We used the NPI *ever* and found no evidence of such facilitation. In fact, we observed a slowdown on reading times in the inference region. Although a failure to find facilitation can sometimes be attributed to a lack of statistical power or to a lack of depth of processing on behalf of the participants, we observed several effects which suggest that such an explanation would be unlikely. First, we found that participants took longer to read the unlicensed NPIs than the licensed NPIs, indicating that the NPI was being processed to some degree. Secondly, participants took longer to read invalid inferences than valid inferences, indicating that they were aware of the polarity context. These secondary findings suggest that if NPIs did facilitate inference making in negative contexts, we would have found facilitation in these experiments.

In Experiment 2, the NPI *ever* scoped over both disjuncts in S2, and it was not repeated in the inference proposition in S3. We conducted another experiment to determine whether these facts played (a) crucial role in our results. This experiment included items where (i) the NPI *ever* was repeated in the inference proposition in S3 and (ii) the NPI *any* occurred inside the second disjunct in S2.

2.3 Experiment 3

Experiment 3 employed the same methodology as Experiment 2, but we made several changes to the items. First, we added the NPI to the third sentence of the *ever* items in the NPI-present condition (a). For example, if the Experiment 2 sentence was

S3 Since/ almost no campers/ have caught a cold,/ the parents are happy,/ and they praise the counselors./

the equivalent Experiment 3 sentence became

S3 Since/ almost no campers/ have ever/ caught a cold,/ the parents are happy,/ and they praise the counselors./

We made this change to test whether the extra reading time associated with the NPI could have been due to participants experiencing difficulty comparing the proposition in the *since* sentence, which did not contain an NPI, against the inference propositions generated in the disjunction sentence.

A second change was that we included modified versions of the *any* items from Experiment 1, the verification task. In Experiment 1, *any* was present in the first disjunct, whereas the inference proposition contained the second disjunct. These were altered so that now *any* was present in the second disjunct of S2, and this same proposition was inferred in S3 (although without repeating the NPI in S3). These were enabled by changing the order of the disjuncts in S2, and replacing the inference proposition in S3 with the now-second disjunct. For example, the ‘local fair’ vignette from Experiment 1 (shown in Table 1) became

S1 After winning a game/ at a local fair,/ children get to choose/ a prize./

S2 Almost no child/ decided on the cotton candy/ or chose any mint chocolates./

S3 Since/ almost no child/ chose mint chocolates,/ the mint chocolates were thrown away,/ and a new prize introduced./

S4 The local fair/ happens/ once a year./

S5 Does the fair happen twice a year?

The NPI *any* was now in the inference proposition, as *ever* was for the *ever* items. Moreover, just as *ever* was not repeated in S3 in Experiment 2, *any* was not repeated in S3 in this experiment.

Finally, we no longer included the unlicensed NPI sentences. We already knew from Experiments 1 and 2 that participants were processing the NPIs and so there was no remaining need for that condition.

2.3.1 Method

Participants: Forty-eight New York University students participated for course credit or payment. They were randomly allocated to one of the three counterbalancing conditions.

Stimuli and design: The items were very similar to those used in Experiment 2. There were two changes, as described in section 1. First, the NPI in the *any* items was now in the inference proposition. Note that this resulted in the inference proposition being different from that used in Experiment 1. For example, in the local fair item, above, ‘decided on the cotton candy’ was the Experiment 1 inference proposition, whereas ‘chose mint chocolates’ was the Experiment 3 inference proposition. We were unable to keep the same inference proposition across experiments because many of the inference propositions used in Experiment 1 were incompatible with *any*. Having changed the inference proposition, we also changed the order of propositions in S2 to ensure that the second disjunct was used as the inference proposition across both experiments.

We also changed the *ever* items by including the NPI in the third sentence of the NPI-present items. For example, item 1 (condition (a)) of Experiment 2 read ‘Since/ almost no campers/ have caught a cold,/ the parents are happy,/ and they praise the counselors./’ whereas the same items and condition in Experiment 3 was ‘Since/ almost no campers/ have *ever*/ caught a cold,/ the parents are happy,/ and they praise the counselors./’ Note that an extra region was required for the Experiment 3 items (the *have ever* region in the previous example). This was because we needed to include the NPI for the NPI-present conditions but not for NPI-absent conditions, and we needed to maintain an identically lengthened inference region. The unlicensed NPI condition was no longer included, but all other aspects of the design and procedure were identical to Experiment 2.

2.3.2 Results—Data were preprocessed using the same criteria as Experiment 2. Five per cent of the items were removed because of incorrectly answered comprehension questions, and a further 5% of responses were removed as outliers.

All the regions of interest were located in S3, since we had no unlicensed NPI condition (c). The principal reason for conducting the experiment was to establish whether changing the scope of *any* would generate increased reading times on the inference region. Figure 3 shows the within-subject differences for the *any* items. On the inference region, participants were marginally slower in the invalid inference condition (d) than the valid condition (b), $t_1(47) = 2.60$, $P = 0.013$, $t_2(23) = 1.55$, $P = 0.13$. They were also significantly slower in the following region: S3.4, $t_1(47) = 3.10$, $P = 0.003$, $t_2(23) = 2.33$, $P = 0.029$. There were no significant differences between the valid and the invalid inference conditions in S3.1 and S3.2, t_1s < 1, indicating that the slowdown is likely to be restricted to the inference region and further regions, although the interaction between region (S3.1, S3.2 and S3.3) and inference validity condition was not significant, $F_1(2,90) = 2.11$, $P = 0.13$. $F_2 < 1$. Thus, there is some evidence that participants were sensitive to the validity of the inference on the *any* items.

The lower line in Figure 3 shows the difference between the NPI-present and NPI-absent valid inference conditions (conditions (b) minus (a)), for the *any* items. If the NPI facilitated processing, the presence of the NPI should reduce reading times and the values on the b-a curve of Figure 3 should be positive. However, mirroring the effect of the NPI on *ever* items in Experiment 2, the NPI slowed down processing on the inference region, $t_1(47) = 2.071$, P

= 0.044, $t_2(23) = 3.034$, $P = 0.006$, and there was no effect prior to the inference region, $t_1(47) = 1.33$, $P = 0.20$, $t_2(23) = 1.76$, $P = 0.093$, ts < 1, for regions S3.1 and S3.2, respectively, although the interaction was only significant in the items analysis, $F_1(2,90) = 2.11$, $P = 0.13$, $F_2(2,46) = 5.70$, $P = 0.006$. There were also significant differences on region S3.4, t1(47) = 2.17, $P = 0.035$, $t_2(23) = 2.33$, $P = 0.013$, suggesting the slowdown in processing overflowed into the next region. In conclusion, allowing *any* to occur in the inference disjunct meant that the NPI slowed down the RTs.

We were also interested in whether the effect on the *ever* items would replicate when we included the NPI in S3. First, we consider the effect of the invalid inference on RTs. The upper line in Figure 4 displays the invalid minus the valid within-subject difference ((d) minus (b)). Note that the inference region is now S3.4 (not S3.3, as it was in the previous experiment) because of the inclusion of the NPI in S3. On the inference region, RTs were significantly slower in the invalid inference condition (d) than in the valid inference condition (b), $t_1(47) = 3.23$, $P = 0.002$, $t_2(23) = 4.04$, $P = 0.001$. This effect was also present in region S3.2, $t_1(47) = 2.22$, $P = 0.031$, $t_2(23) = 2.42$, $P = 0.024$, and marginally so in S3.1 $t_1(47) = 2.21$, $P = 0.032$, $t_2 = 1.96$, $P = 0.062$, but not in S3.3, $t_1(47) = 1.39$, $P = 0.17$, $t_2(23) = 1.47$, $P = 0.16$. When regions S3.1 to S3.4 were analysed together, a main effect of inference validity was observed, $F_1(1,135) = 16.66$, $P < 0.0005$, $F_2(1,23) = 16.07$, $P = 0.001$, and there was some evidence of an interaction, $F_1(3,135) = 1.35$, $P = 0.26$, $F_2(3,69) = 3.34$, $P = 0.024$.

The effects of the NPI can be seen in the lower curve of Figure 4. Note that region S3.3 has an extra word (the NPI) in condition (a), making interpretation of RT differences complicated in this region. On the inference region, sentences that included the NPI were read more slowly than sentences that did not, although the effect was only significant in the participants analysis, $t_1(47) = 2.39$, $P = 0.021$, $t_2(23) = 1.45$, $P = 0.16$. Similar effects were present in regions prior to the inference region, however, where a main effect of NPI presence was observed (S3.1, S3.2 and S3.4, but not including S3.3 because it contains an extra word), $F_1(1,45) = 10.11$, $P = 0.003$, $F_2(1,23) = 5.73$, $P = 0.025$, but there was no evidence of an interaction, Fs < 1. RTs were significantly longer in the NPI-present conditions for S3.1 and S3.2, $t_1(47) > 2.04$, $P < 0.05$, $t_2(23) > 2.10$, $P < 0.05$. Thus, we observed reading times that were significantly longer in the NPI-present sentences than in the NPI absent sentences, as we did in Experiment 2. Finally, we consider the effects of the NPI across both sets of items. The inference region for the *any* items was S3.3, whereas for the *ever* items it was S3.4. We therefore used these scores as the dependent measure in an ANOVA analysis with NPI presence and NPI type as factors. The overall effect of the NPI presence was to slowdown reading time on the inference region, $F_1(1,45) = 9.75$, $P = 0.003$, $F_2(1,46) = 9.24$, $P = 0.004$, and there was no significant interaction between NPI types and NPI presence, $F_s < 1$. The effect of NPI presence was therefore equal across NPI types. Furthermore, comparing regions S3.1, S3.2 and S3.3 for the *any* items and S3.1, S3.2 and S3.4 for the *ever* items revealed that the effect of NPI presence was greatest on the inference region, as shown by an interaction between region and NPI presence, $F_1(2,90) = 3.75$, $P = 0.027$, $F_2(1,46) = 3.74$, $P = 0.028$, but no NPI presence by NPI type interaction, $F_1(1,45) = 1.70$, $P = 0.20$, $F_2(1,46) = 1.19$, $P = 0.28$, nor region by NPI presence by NPI type interaction, $F_1(2,90) = 2.35$, $P = 0.11$. $F_2(2,92) = 2.94$, $P = 0.058$.

2.3.3 Discussion—In this experiment, we included items with *any* in the second disjunct, so that *any* modified the inference proposition. We found that the presence of the NPI slowed reading times on S3, particularly on the inference region, as it did for the *ever* items in Experiment 2. Clearly, there was no evidence that the NPI facilitated processing of the inference, contrary to the association account of NPI licensing.

We also altered the *ever* items by including the NPI in the third sentence for the NPI-present items. We were concerned that the effects of the NPI were due to the difficulty participants might have experienced comparing the propositions generated with the NPI in the second sentence against propositions presented without the NPI in the third. In contrast to this hypothesis, we observed significantly slower reading times on the inference region in the NPI-present condition than the NPI absent condition, and when we considered the *any* and the *ever* items together, there was a significant slowdown on the inference region and no significant difference between the two quantifiers. Thus, the effects of the NPI on the inference in this experiment, in which we included the NPI in the *since* sentence, were similar to those in Experiment 2, in which no NPI occurred in the *since* sentence. However, the effect of the NPI was less robust than previously (the slowdown on the inference region was significant only in the participants analysis) and spillover effects from the extra word in the NPI condition could have contributed to the slowdown. (This was not an issue in the previous experiment, nor for the *any* items, because the inference sentences were identical in all relevant conditions.) We cannot therefore completely rule out the possibility that differences between inferences derived from the *since* sentence and propositions presented in the third sentence contributed something to the observed slowdown. This issue is considered further in section 3.

Finally, we observed significantly slower reading times as a result of the NPI in regions prior to the inference region in S3, as we did in Experiment 2. It is therefore possible that integrating NPI sentences into the discourse is more complex than integrating sentences without an NPI. However, we did not observe this effect with the *any* items and so the effect appears to be NPI specific or related to the items that supported *ever*. Moreover, the effects on the inference region were significantly greater than those on prior regions in Experiments 2 and 3 (at least when *ever* and *any* items are combined), as measured by the significant interaction terms. Thus, at least some of the slowdown we observe on the inference region cannot be due to whatever it is that causes the slowdown on the earlier regions. We leave the locus of this effect to future research.

3 GENERAL DISCUSSION

The goal of this study was to establish whether the processor recognizes the relationship between decreasingness and NPI licensing. We argued that a domain-widening plus proposition-strengthening theory would plausibly predict that the presence of a licensed NPI facilitates inferences from sets to subsets (the association model). To this end, we constructed vignettes that contained sentences with either a decreasing or a non-decreasing quantifier followed by inferences that were valid or invalid, respectively. Crucially, we manipulated whether these texts contained an NPI. According to our construal of the association model, the presence of the NPI should have facilitated processing by making it more likely that participants would evaluate the inference correctly or by speeding up the interpretation of the inference. Contrary to this hypothesis, however, we saw no evidence that the presence of the NPI facilitated processing, whether participants were given explicit reasoning tasks (Experiment 1) or whether they read sentences in a self-paced reading task (Experiments 2 and 3), and therefore we found no evidence supporting the claim that decreasingness and NPI licensing are linked. It is unlikely that our results were simply due to low sensitivity in general because we found three other reliable effects involving the NPI and the inferences: (i) participants read the NPI more slowly when it was unlicensed than when it was licensed, (ii) invalid inferences were read more slowly than valid inferences and (iii) the presence of the NPI made processing reliably more difficult on the inference region (i.e. the results were significantly in the opposite direction compared to the facilitation predictions).

The results are consistent with the hypothesis that the processor does not recognize the relationship between NPIs and inferences to subsets, i.e. the basic tenet of the dissociation model. Before accepting this conclusion, however, we consider other potential explanations for our results. First, suppose that the processor does recognize the relationship between NPIs and inferences, but that any facilitatory effects were obscured in the contexts that we tested. For example, the quantifier at the beginning of the disjunction sentences might have signalled the DE context sufficiently strongly that the licensed NPI was not necessary for the grammar to fully recognize the context. In effect, the context could have been maximally recognized by the grammar before encountering the NPI. The results of Experiment 1 argue against this explanation, however. If participants could so easily identify the context on the basis of the quantifier alone, performance in the inference task should have been very high without the NPI, because the validity of the inference can be perfectly predicted from the context. Instead, participants correctly rejected only 58% of the invalid inferences and accepted only 87% of the valid inferences. Although this is not unusually poor performance compared to other reasoning tasks reported in the literature, participants clearly did not find it especially easy to identify DE contexts without the NPI.

A second possible explanation for why we did not find facilitating effects of the NPI is that the context signalling effect may not have been sufficiently long lasting to carry over from the second sentence to the third sentence. This account assumes that the marker for decreasingness or the effect of the NPI fades across time (or linguistic input). A strong argument against this, however, is that the effects of the NPI were sufficiently long lasting that they slowed down reading in S3 when the NPI was present in the proposition (in S2). It would be unlikely that the NPI would have this effect while not continuing to signal decreasingness (if it signals it at all). We conclude that these two, linguistically uninteresting, explanations for the absence of any facilitation effects are therefore implausible.

While writing up this article, we learned from E. Chemla (personal communication) that he had conducted a pilot study involving inference verification in the presence/absence of an NPI in French and also found no facilitatory effect, converging with the results of our own Experiment 1. In Chemla's experiment, the premise was read independently and did not contribute to response time. The context was set up so as to exclude the one specific European interpretation. The stimuli involved the quantifiers *aucun* 'no', *moins de 4* 'less than 4', *plupart … ne … pas* 'most … didn't' and *plus de 4 … ne … pas* 'more than 4 … didn't', and were of the following shape:

(15) Aucun chien n'a touché un / le moindre européen(s).

'No dog touched a / a single European'

$\overset{?}{\Rightarrow}$

Aucun chien n'a touché de français.

'No dog touched any Frenchman'

(16) La plupart des chiens n'ont pas touché un / le moindre européen(s).

'Most of the dogs didn't touch a / a single European'

$\overset{?}{\Rightarrow}$

La plupart des chiens n'ont pas touché de français.

'Most of the dogs didn't touch any Frenchman'

Both the percentage of correct answers and the mean response times for correct answers were worse when the premise contained the NPI *le moindre* than when it contained indefinites with *un*. Chemla's results, while very preliminary, suggest that the lack of facilitation in our Experiment 1 is not explained by some accidental feature of the design.

Instead of a facilitation effect of the NPI, in the reading time experiments we found that introducing an NPI slowed down inference processing. When the inference verification involved a proposition that contained an NPI (in the preceding sentence), reading times were longer compared to when no NPI was present. For example, 'Since almost no campers have caught a cold' took longer to read when the preceding sentence was 'Almost no campers have ever had a sunburn or caught a cold' than when the preceding sentence did not contain *ever*. Furthermore, in Experiment 2, the slowdown was localized to the inference region (*caught a cold*) and the region immediately following it. Thus, the slowdown was directly related to inference verification, and not to general complexity issues associated with processing NPIs.

We formulate two hypotheses consistent with the reported patterns across experiments:

(17) No facilitation plus somewhat costly NPI processing:

> NPI presence does not improve either the accuracy or the speed of inference processing. On the other hand it incurs some cost that is manifested in increased reading times.

(18) Some facilitation plus very costly NPI processing:

> NPI presence does facilitate inference processing in some way and to some extent, but it also incurs a cost that is large enough both to wipe out all facilitatory effects and to additionally increase reading times.

In what way can the processing of the NPI be costly? One possibility is that recognizing the relationship between the propositions with *ever/any* (in S2) and those without *ever/any* (in S3) was difficult. But it is unlikely that absence of the NPI from S3—i.e. a 'mismatch effect' —is the main factor in the slowdown. In Experiment 2, S3 always contained the second disjunct of S2 and therefore S3 lacked a big chunk of S2, not just the NPI. For example:

S2 Almost no campers have **ever** had a sunburn or caught a cold.

S3 Since almost no campers have caught a cold,

S2 At most half of the plants have **ever** died or lost leaves.

S3 Since at most half of the plants have lost leaves,

Since S3 was never a verbatim repetition of S2, the stimuli did not specifically generate an expectation for the NPI to occur in S2. A more complex form of the mismatch argument may be that participants required extra time to recognize the relationship between the proposition with the NPI in S2 and the proposition without the NPI in S3. But note that the key idea of the scalar account is that the NPI is licensed if and only if the proposition with the NPI entails its counterpart without the NPI. In other words, the diagnostic of NPI licensing involves a mental comparison of 'mismatching propositions'. If the scalar account is correct, participants who registered that the NPI in S2 was licensed should have already efficiently compared S2 and S3. Therefore, the possibility that the need to compare propositions with and without NPIs determined the reading times in S3 is more compatible with the No Facilitation hypothesis in (17) than with the Some Facilitation hypothesis in (18). This could be further tested by experiments focusing on the repetition v. omission of the NPI in the inference proposition. Finally, the NPI was repeated in Experiment 3 and some slowdown nevertheless occurred, whereas the mismatch explanation would predict a complete elimination of the slowdown.

An alternative possibility is that the semantics/pragmatics of the NPI incurs a significant processing cost. All three theoretical accounts reviewed in section 1 could predict this. On the particular version of the non-veridicality account presented in Giannakidou (1998), some NPIs are referentially deficient: they contain non-deictic variables and thus have to be bound by, or be anaphoric to, an antecedent (for a purely syntactic version of this idea see Progovac 1994). Thus, their processing cost should be similar to that of bound or anaphoric pronouns. On the Ladusaw-de Swart&Sag-Postal-Szabolcsi account, the factoring out of the negative component of the NPI's lexical representation and the formation of a polyadic quantifier with the negation component of the licensor may well be costly. On the Kadmon&Landman-Krifka-Lahiri-Chierchia account, the NPI itself induces scalar implicatures. This aspect of the account was not detailed in the section 1; we summarize it here, based on Chierchia (2006: 554–60). Chierchia follows Krifka and Lahiri in attributing an *even*-like flavour to the base meaning of the NPI *any*. 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 *any* true. (This implicature can only be true in a decreasing local environment; in such an environment, the *any*-sentence will entail its counterpart with a plain indefinite, which always quantifies over some particular domain.) Departing from Grice, implicatures are added and strengthened meanings are calculated recursively, at every step of the sentence's composition. Domain widening and implicature calculation are plausibly costly real-time operations.

The upshot is that all the recent theoretical accounts are in principle capable of explaining the findings (no observable facilitation of accuracy, some slowdown in reading times). What our findings clearly rule out is an account that predicts that NPIs should have a squarely facilitatory effect, as in Dowty (1994). Further work might tease apart the localization and magnitude of the effects, and determine whether one of the models is favoured by the processing data.[7]

As one issue of interest, notice that while especially stressed NPIs undeniably have an *even*-flavoured meaning, not all NPIs do. Some examples of NPIs without domain widening are n-words interpreted as NPIs (as is observed in Chierchia 2006), occurrences of unstressed *any* applied to rigorously defined domains ('The empty set does not have *any* proper subsets' is fully acceptable, but it does not mean 'even a marginal proper subset', Krifka 1995), and items like the adverb *anymore* ('He doesn't live here *any*more' 'He lived here and that has changed'), the auxiliary *need* ('He *need* not come early') and others (van der Wal 1999). The existence of such NPIs is one reason why some accounts maintain that the phenomenon of NPI licensing *per se* is not an essentially scalar matter. On the other hand, the basically non-scalar theories may freely acknowledge that some NPIs do have an *even*-flavour that has to be taken into account in the full description of their distribution and meaning (Szabolcsi 2004, and especially Giannakidou 2007). Such representatives of the dissociation model predict that the processing of an NPI is more costly when it actually carries scalar implicatures.

Chierchia's (2006) account accommodates the existence of NPIs without actual domain widening in the following way. In contrast to items like *some* and *many*, whose scalar alternatives can be deactivated and thus their implicatures ('but not all') suspended in appropriate contexts, *any* is grammaticized to always activate a set of domain alternatives. On the other hand, Chierchia requires the proposition with the widest domain of quantification only to entail its counterparts with particular domains; that is, it has to be either stronger than or equivalent to them. ‘Domain widening, as implemented here, is a *potential* for domain widening’ (Chierchia 2006: 559, emphasis in the original). In this way, his account does not distort interpretations. However, the combined effect of the grammaticized activation of domain alternatives and the recursive computation of scalar implicatures is that NPIs will incur the same processing cost regardless of whether they actually involve domain widening (‘The camper has not suffered *ANY* bruises’) or not (‘The empty set does not have *any* proper subsets’). This prediction contrasts with that of the dissociation model.

---

[7] Among other things, it would be very interesting to find out whether there is a slowdown even if the NPI occurs in a segment of the preceding sentence that is not part of the inference proposition. In fact, the *any* items mentioned in note 6 were intended to serve that purpose but, as was pointed out in that footnote, for some reason that batch of items yielded non-significant results in general.

Further work may be able to determine which prediction is borne out by processing. The question is whether domain alternatives are always computed or whether they are only computed when the NPI involves actual domain widening. In this regard, the hypothesis parallels that of processing work on scalar implicatures (e.g. Noveck & Posada 2003; Bott & Noveck 2004; Breheny *et al*. 2006). These researchers have contrasted neo-Gricean accounts of how scalar implicatures are processed (e.g. Levinson 2000; Chierchia 2004) with context-dependent theories (e.g. Sperber &Wilson 1985). The crucial difference between these accounts is that neo-Gricean accounts predict scalar implicatures to be computed by default on encountering a scalar term like *some*, that is, the scalar alternatives are always calculated and scalar implicatures (denial of the stronger element in the scale) go through unless cancelled. The context-dependent account predicts that the scalar alternatives are only calculated in specific contexts, that is, there is no default computation of scalar alternatives. The results of these processing investigations have been that interpreting sentences with scalar implicatures, like ‘Some [but not all] children are in the classroom’ requires more processing time than interpreting the sentences without the implicature, as in ‘Some [and possibly all] of the children are in the classroom’ (e.g. Bott & Noveck 2004), thus arguing against a strict default theory.

Somewhat modifying his 2004 proposal, Chierchia (2006) assumes that the default activation of alternatives can be suspended to begin with. Scalar terms have a strong [$+\sigma$] variant with alternatives, and a weak [$-\sigma$] variant without alternatives. The weak variant is employed in those cases that on the previous account involved implicature cancellation. The claim that NPIs always have active domain alternatives is technically expressed as they having only strong variants in Chierchia (2006). Therefore, it is significant that earlier work has demonstrated that processing with and without implicatures incurs different costs. Similar questions to what have been asked about non-NPI scalar terms in the experimental studies cited can now be asked about NPIs, contrasting cases of bona fide domain widening with cases where the NPI either lexically or contextually fails to actually widen the domain. Do both slow down the processing of decreasing inferences, or do only bona fide domain-widening NPIs do so?[8]

4 CONCLUSION

The classical explanation for how NPIs are licensed is that they are allowable only in the scope of a downward entailing operator. We argued that a plausible processing model derived from such an account would predict that the presence of an NPI should facilitate the processing of inferences from sets to subsets: the NPI should highlight the decreasingness of the context. Yet, we did not observe the expected facilitation effects. In Experiment 1, we found that the presence of an NPI had no effect whatsoever on the likelihood of participants correctly verifying inferences, while in Experiments 2 and 3 we observed the reverse finding; that the NPI significantly slowed participants' inference verification strategies. While we cannot categorically demonstrate which of several explanations is responsible for this slowdown, all the explanations invoke the NPI in the inference verification process, so it is difficult to argue that the NPI was not a significant contributor to the verification process in general.

---

[8] Not all scalar theories of negative polarity are purely domain-widening theories, see e.g. Krifka (1995), so the predictions will have to be tested in a differentiated manner.

Our findings suggest that NPIs do not play an important facilitatory role in inference making. The most straightforward implication is that the processor does not recognize the relationship between NPIs and decreasingness (at least not in the way in which the simple association model predicts it should do). The challenge for future researchers is to address what role the NPI plays in the inference-making process, and, more generally, which of the models that we outlined in section 3 is the most accurate processing model of NPI licensing. We have indicated various experiments that could be conducted to this end.

APPENDIX

EXPERIMENTAL STIMULI

The list below exemplifies each type of stimulus with one entry; the experiments used two entries of each type. The full list can be obtained from the authors. Items 1–12 all contain *ever*, using the valid inference version with a DE operator. These items were used in all three experiments. Below they appear as presented to the participants using a self-paced reading paradigm in Experiment 2. The slashes separate the regions. Experiment 3 was similar except that the NPI was repeated in the *since* sentence (S3), in addition to appearing in the disjunction sentence (S2). In Experiment 1, S3 was not a declarative with *since* but a question using the 'Would it be reasonable …?' construction discussed in the text.

1. Our camp/ is on/ Staten Island./

   Almost no campers/ have ever had/ a sunburn/ or caught a cold./

   Since/ almost no campers/ have caught a cold,/ the parents are happy,/ and they praise the counselors./

   They will/ use the camp/ again next year./

   Is the camp in Staten Island?

2. I keep/ my collectibles/ on the counter./

   No more than five pieces/ have ever fallen/ or gotten smashed./

   Since/ no more than five pieces/ have gotten smashed,/ the arrangement seems safe,/ and I will stick with it./

   I may/ add/ new pieces./

   Do I keep my collectibles on the counter?

3. The club/ hikes in/ the Palisades./

   At most five members/ have ever spotted/ deer/ or found bluebells./

   Since/ at most five members/ have found bluebells,/ the women are grumbling,/ and they are considering other clubs./

   The kids/ don't care./

   Is this a hiking club?

4. We assemble/ our own/ furniture./

   Not many chairs/ have ever had/ wobbly legs/ or fallen apart./

   Since/ not many chairs/ have fallen apart,/ we'll get some tables,/ and perhaps more complicated pieces./

   One saves/ on these/ purchases./

   Do we assemble furniture?

5. Mary got/ a toy set/ from her grandma./

   Less than five pieces/ have ever been/ damaged/ or gotten lost./

   Since/ less than five pieces/ have gotten lost,/ the set is in good shape,/ and Mary will save it for her children./

   She only/ keeps/ quality stuff./

   Did Mary get a toy set from her grandma?

6. I leave/ messages about/ my theater group./

   Very few people/ have ever come/ to the show/ or called back./

   Since/ very few people/ have called back,/ my group is unhappy,/ and they say I am lazy./

   They want/ to place calls/ themselves./

   Is my group unhappy?

7. One of/ our friends/ is sick./

   Almost nobody/ has ever sent/ flowers or written a card./

   Since/ almost nobody/ has written a card,/ our friend feels neglected,/ and she is complaining./

   She will/ return to work/ next month./

   Is our friend sick?

8. Every summer/ Tom grows/ tomatoes./

   At most half the plants/ have ever died/ or lost leaves./

   Since/ at most half the plants/ have lost leaves,/ Tom is quite satisfied,/ and he doesn't consider growing carrots./

   He likes/ his garden./

   Does Tom grow tomatoes?

9. I keep/ an eye on/ the enrollments./

   No more than fifty people/ have ever signed up/ for squash/ or taken karate./

   Since/ no more than fifty people/ have taken karate,/ some classes are cancelled,/ and the program will move to the second floor./

   The coaches/ are holding/ a meeting./

   Are the coaches holding a meeting?

10. Judy likes/ to display the books/ she gets/ for presents./

Not many of the books/ have ever been/ thick/ or had two volumes./

Since/ not many of the books/ have had two volumes,/ the display is small,/ and it fits on her shelf./

The shelf/ is above/ Judy's desk./

Is Judy getting books for presents?

11. I use/ a lab/ computer./

Less than five of my files/ have ever been/ tampered with/ or gotten infected./

Since/ less than five of my files/ have gotten infected,/

I like the arrangement,/ and I trust the other users./

I can't afford/ a laptop/ now./

Do I use lab computer?

12. Max commutes/ from/ Tempe./

Very few of his flights/ have ever been/ overbooked/ or suddenly canceled./

Since/ very few of his flights/ have been suddenly canceled,/

Max's commute is easy,/ and he doesn't worry about it./

Max likes/ to live/ in Tempe./

Does Max work outside his home town?

Items 13–24 all contain *any*. The results for these items were presented in Experiments 1 and 3. The items are in the form presented to participants in Experiment 3 (the disjuncts should be reversed for Experiment 1).

13 Julie keeps/ tropical fish and plants/ in her aquarium./

No more than five fish/ have had a disease /or have eaten any plants.

Since/ no more than five fish/ have eaten plants,/ Julie is quite pleased/ and may buy more fish./

She will buy/ some new/ plants too./

Will Julie get some new plants?

14 The ice cream shop/ around the corner/ sold a lot of/ sundaes this weekend./

Not many people/ chose low calorie sundaes /or bought any fruit sorbets.

Since/ not many people/ bought the fruit sorbets/ they won't offer them again,/ and will order ice cream instead./

Next weekend/ they hope to make/ more money./

Did the ice cream shop sell a lot of sundaes?

15 This department store/ had a/ housing goods sale./

Less than five salespeople/ sold living room rugs /or made any significant profit.

Since/ less than five people/ made significant profit,/ the store lost money/ and they will lose their jobs./

The store/ usually has sales/ in the fall/ and the spring./

Did the store have a housing goods sale?

16 Our family/ always cooks/ a traditional meal/ for Thanksgiving./

Very few people/ take a second helping of sweet potatoes /or eat any creamed onions.

Since/ very few people/ eat creamed onions,/ there are some leftovers/ and we have them again the next day./

There is usually/ turkey left over/ too./

Does the family cook a traditional meal for Thanksgiving?

17 I went/ to the new/ train station yesterday/ to ask about/ my trip to Nantucket./

Almost nobody/ was buying tickets /or boarding any trains.

Since/ almost nobody/ was boarding trains,/ the lines were short/ and I could ask about my train times./

I want/ to leave/ as soon as possible./

Did I go to the train station?

18 The Atlantic Hotel/ sends its napkins/ to be washed/ on Wednesdays./

No more than fifty napkins/ are spotted with bleach /or come back with any stains.

Since/ no more than fifty napkins/ come back with stains,/ the napkins are reused,/ and the hotel buys new ones the next season./

The hotel/ is especially busy/ during the summer./

Is the hotel busy in the summer?

19 Paradise Cruise Lines/ had a holiday cruise/ that lasted a week./

Less than five of the passengers/ felt seasick during the cruise /or had any serious complaints.

Since/ less than five of the passengers/ had serious complaints,/ the captain was happy,/ and he enjoyed the cruise./

The captain/ hopes the next trip/ will go/ as smoothly./

Did the cruise last a week?

20 The local supermarket/ conducted a survey/ of its customers./

Very few of the surveys/ were returned late /or had any blank responses.

Since/ very few of the surveys/ had blank responses,/ the results were tallied/ and the information was passed on to the managers./

The supermarket/ conducted a survey/ again the next year./

Was it a supermarket that conducted the survey?

21 After winning a game/ at a local fair,/ children get to choose/ a prize./

Almost no child/ decided on the cotton candy/ or chose any mint chocolates./

Since/ almost no child/ chose mint chocolates,/ the mint chocolates were thrown away,/ and a new prize introduced./

The local fair/ happens/ once a year./

Does the fair happen twice a year?

22 Children from/ a boys and girls club/ went on a campaign/ to raise money./

At most five boys/ got a large donation/ or sold any chocolate candy./

Since/ at most five boys/ sold chocolate candy,/ the club will campaign again,/ and go to a different neighborhood./

Selling chocolate/ is a good way/ to raise money./

Were the children from a basketball team?

23 A clothing designer/ asked a focus group/ to pick between buttons/ of various shapes and colors./

At most half the people/ picked square buttons /or chose any purple buttons.

Since/ at most half the people/ picked purple buttons,/ the results were discouraging,/ and the color was discontinued./

There were/ four different/ button colors./

Were there ten different button colors?

24 The curator/ of an art gallery/ set up a new exhibit/ of paintings and sculptures./

Not many of the artists/ sculpted large pieces/ or painted any big pictures.

Since/ not many of the artists/ painted big pictures,/ the painting area was small,/ and left lots of space for the sculptures./

Setting up/ an art exhibit/ is itself/ an art./

Did the exhibit feature photography?

Acknowledgments

Lewis Bott was supported by NIMH Grant MH41704 awarded to Gregory L. Murphy for part of this work. Further support was provided by NIH grant R01HD056200 awarded to Brian McElree. We are grateful to Lyn Frazier, Emmanuel Chemla and Mark Steedman for helpful discussion of an earlier version of the manuscript, to the editor and two reviewers for comments, and to Gregory L. Murphy for his advice and the use of his laboratory. We also thank Jessica Piercy for collecting the data and Tuuli Adams and Peter Liem for help with writing the items.

References

Altman A, Peterzil Y, Winter Y. Scope dominance with upward monotone quantifiers. Journal of Logic, Language and Information. 2005; 14:445–55.

Bernardi, R. Ph.D dissertation. Utrecht, The Netherlands: 2002. Reasoning with Polarity in Categorial Type Logic.

Bott L, Noveck I. Some utterances are underinformative: the onset and time course of scalar inference. Journal of Memory and Language. 2004; 51:437–57.

Breheny R, Katsos N, Williams J. Are generalised scalar implicatures generated by default? An online investigation into the role of context in generating pragmatic inferences. Cognition. 2006; 100:434–63. [PubMed: 16115617]

Chierchia, G. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In: Belletti, A., editor. Structures and Beyond. Oxford University Press; Oxford: 2004.

Chierchia G. Broaden your views: implicatures of domain widening and the “logicality” of language. Linguistic Inquiry. 2006; 37:535–91.

De Decker, P.; Larsson, E.; Martin, A. Workshop on Polarity from Different Perspectives. New York University; 2005. Polarity judgments: an empirical view. http://www.nyu.edu/gsas/dept/lingu/events/polarity/posters/dedecker-larsson-martin.pdf

den Dikken, M. Parasitism, secondary triggering, and depth of embedding. In: Zanuttini, R.; Campos, H.; Herburger, E.; Portner, PH., editors. Cross-linguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture. Georgetown University Press; Washington, DC: 2006. p. 151-75.

de Swart H, Sag IA. Negation and negative concord in Romance. Linguistics and Philosophy. 2002; 25:373–417.

Dowty, D. The role of negative polarity and concord marking in natural language reasoning. In: Harvey, M.; Santelman, L., editors. Proceedings from Semantics and Linguistic Theory IV. Cornell University; Ithaca: 1994. p. 114-45.

Ferreira F, Patson N. The “good enough” approach to language comprehension. Language and Linguistics Compass. 2007; 1:71–83.

Fyodorov Y, Winter Y, Francez N. Order-based inference in “Natural Logic”. Logic Journal of the IGPL. 2003; 11:385–416.

Gajewski J. Licensing strong NPIs. Proceedings of the 31st Annual Penn Linguistics Colloquium. Papers in Linguistics. 2008; 14(1):163–76.

Geurts B. Reasoning with quantifiers. Cognition. 2003a; 86:223–51. [PubMed: 12485739]

Geurts B. Monotonicity and syllogistic inference: a reply to Newstead. Cognition. 2003b; 90:201–4.

Geurts B, van der Slik F. Monotonicity and processing load. Journal of Semantics. 2005; 22:97–117.

Giannakidou, A. Polarity Sensitivity as (Non)Veridical Dependency. Linguistik Aktuell/Linguistics Today 23. John Benjamins; 1998.

Giannakidou A. *Only*, emotive factives, and the dual nature of polarity dependency. Language. 2006; 82:575–603.

Giannakidou A. The landscape of EVEN. Natural Language and Linguistic Theory. 2007; 25:39–81.

Heim, I. Tancredi, C., et al., editors. Little. Proceedings of Semantics and Linguistic Theory. 2006. http://research.nii.ac.jp/salt16/proceedings.html

Kadmon N, Landman F. Any. Linguistics and Philosophy. 1993; 16:353–422.

Koller, A.; Niehren, J. ESSLLI Reader; 1999. Scope underspecification and processing. http://www.ps.unisb.de/Papers/abstracts/./ESSLLI:99.ps

Krifka M. The semantics and pragmatics of polarity items. Linguistic Analysis. 1995; 25:209–57.

Ladusaw, W. Polarity Sensitivity as Inherent Scope Relations. Garland, New York: 1980.

Ladusaw, W. Expressing negation. In: Barker, C.; Dowty, D., editors. Proceedings of Semantics and Linguistic Theory II. OSU; Columbus, OH: 1992. p. 237-61.

Lahiri U. Focus and negative polarity in Hindi. Natural Language Semantics. 1997; 6:57–123.

Levinson, S. Presumptive Meanings. The MIT Press; Cambridge, MA: 2000.

McKoon G, Ratcliff R. Inference during reading. Psychological Review. 1992; 99:440–66. [PubMed: 1502273]

Newstead SE. Can natural language semantics explain syllogistic reasoning? Cognition. 2003; 90:193–9. [PubMed: 14599753]

Noveck I, Posada A. Characterising the time course of an implicature. Brain and Language. 2003; 85:203–10. [PubMed: 12735938]

Paterson KB, Sanford AJ, Moxey LM, Dawydiak EJ. Quantifier polarity and referential focus during reading. Journal of Memory and Language. 1998; 39:290–306.

Postal, PM. Workshop on Polarity from Different Perspectives. New York University: 2005. Suppose (if only for an hour) that NPIs are negation-containing phrases. http://www.nyu.edu/gsas/dept/lingu/events/polarity/papers/postal-paper.pdf

Progovac, L. Cambridge Studies in Linguistics. Cambridge U.P.; Cambridge: 1994. Negative and Positive Polarity: A Binding Approach.

Saánchez-Valencia, V. Ph.D dissertation. Amsterdam, The Netherlands: 1991. Studies on Natural Logic and Categorial Grammar.

Simons, M. Review of Giannakidou, Polarity Sensitivity as Non Veridical Dependency. 1999. http://linguistlist.org/issues/10/10-1152.html

Sperber, D.; Wilson, D. Relevance: Communication and Cognition. Blackwell; Oxford: 1985.

Szabolcsi A. Positive polarity–negative polarity. Natural Language and Linguistic Theory. 2004; 22:409–52.

Szabolcsi, A. Scope and binding. In: Maienborn, C.; von Heusinger, K.; Portner, P., editors. Semantics: An International Handbook of Natural Language Meaning. Mouton de Gruyter; Berlin-NewYork: 2006.

Tunstall, S. Ph.D. thesis. University of Massachusetts; Amherst, MA: 1998. The Interpretation of Quantifiers: Semantics and Processing.

Verbrugge, S.; Schaeken, W. The ballad of if and since: never the twain shall meet?. Proceedings of the 28th Annual Conference of the Cognitive Science Society; Vancouver, Canada. 26–29 July 2006; 2006. p. 2305-10.http://www.cogsci.rpi.edu/csjarchive/Proceedings/2006/docs/p2305.pdf

van Benthem, J. Language in Action. North-Holland; Amsterdam, The Netherlands: 1991.

van der Wal, S. Ph.D dissertation. Groningen, The Netherlands: 1999. Negative Polarity Items and Licensing: Tandem Acquisition.

von Fintel K. NPI-licensing, Strawson-entailment, and context dependency. Journal of Semantics. 1999; 16:97–148.

Zwarts F. Negatief polaire uitdrukkingen 1. GLOT. 1981; 4:35–132.

Zwarts F. Nonveridical contexts. Linguistic Analysis. 1995; 25:286–312.

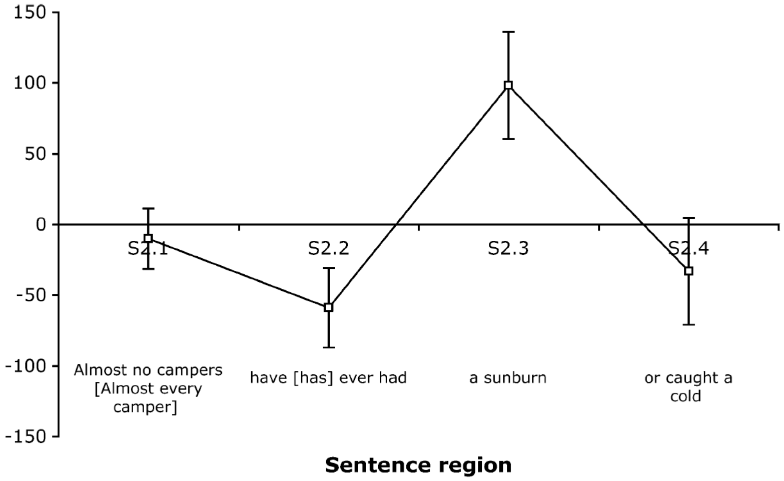

Figure 1.

Reading time differences (ms) for S2, Experiment 2. NPI present, unlicensed (c) minus licensed (a). Values on the *x*-axis correspond to different regions of the sentence. The appropriate section of the ‘Campers’ item is shown on each region (words in square brackets correspond to the (c) version). Positive difference scores indicate longer reading times in the unlicensed NPI condition. Error bars correspond to the standard error of the difference for each region (subject analysis).

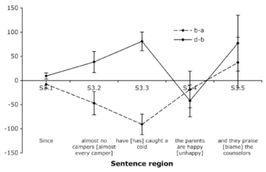

Figure 2.

Reading time differences (ms) for S3, Experiment 2. The upper line is the invalid (d) minus valid (b) inference conditions, no NPIs. The lower line is the valid inference, NPI absent (b) minus valid inference, NPI present (a) condition. Values on the *x*-axis correspond to different regions of the sentence. The appropriate section of the 'Campers' item is shown on each region (words in square brackets correspond to the (d) version). Error bars correspond to the standard error of the difference for each region (subject analysis).

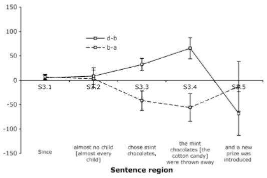

Figure 3.

Reading time differences (ms) for the *any* items, S3, Experiment 3. The upper line is the invalid (d) minus valid (b) inference conditions, no NPIs. The lower line is the valid inference, NPI absent (b) minus valid inference, NPI present (a) condition. Values on the *x*-axis correspond to different regions of the sentence. The appropriate section of the ‘local fair’ item is shown on each region (words in square brackets correspond to the (d) version). Error bars correspond to the standard error of the difference for each region (subject analysis).

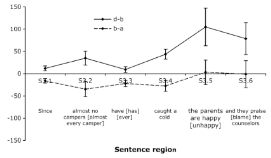

Figure 4.

Reading time differences (ms) for the *ever* items, S3, Experiment 3. The upper line is the invalid (d) minus valid (b) inference conditions, no NPIs. The lower line is the valid inference, NPI absent (b) minus valid inference, NPI present (a) condition. Values on the *x*-axis correspond to different regions of the sentence. The appropriate section of the ‘Campers’ item is shown on each region (words in square brackets correspond to the (a) or (d) version). Error bars correspond to the standard error of the difference for each region (subject analysis).

Table 1

Examples of the stimuli used in Experiment 1

	***Ever* example**	***Any* example**
Context sentence	Our camp is on Staten Island.	After winning a game at a local fair, children get to choose a prize.
(a) decreasing with NPI	Almost no campers have ever had a sunburn or caught a cold.	Almost no child chose any mint chocolates or decided on the cotton candy.
(b) decreasing without NPI	Almost no campers have had a sunburn or caught a cold.	Almost no child chose mint chocolates or decided on the cotton candy.
(c) non-decreasing with NPI	Almost every camper has ever had a sunburn or caught a cold.	Almost every child chose any mint chocolates or decided on the cotton candy.
(d) non-decreasing without NPI	Almost every camper has had a sunburn or caught a cold.	Almost every child chose mint chocolates or decided on the cotton candy.
Inference question	Would it be reasonable to say that almost no [almost every] camper[s] have [has] caught a cold?	Would it be reasonable to say that almost no [almost every] child decided on the cotton candy?

Table 2

Quantifier pairs

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

Table 3

Proportion ‘valid’ responses for Experiment 1. Mean proportions ‘valid’ responses as a function of inference type (valid or invalid), presence of NPI (+ for present, − for absent) and NPI type (*ever* or *any*). Standard deviations (SDs) in parentheses. Proportions correct for (a) and (b) are equal to the tabulated value, but proportions correct for (c) and (d) are equal to 1 minus the tabulated value

	(a)Valid, NPI+	**(b)Valid, NPI−**	**(c)Invalid, NPI+**	**(d)Invalid, NPI−**
any	0.89 (0.17)	0.87 (0.14)	0.51 (0.31)	0.38 (0.31)
ever	0.85 (0.18)	0.87 (0.20)	0.46 (0.31)	0.44 (0.37)
Total	0.87 (0.13)	0.87 (0.14)	0.49 (0.29)	0.42 (0.33)

Table 4

Staten Island vignette

Sentence	Conditions (a) [NPI+] and (b)	Conditions (c) [NPI+] and (d)
S1	Our camp/ is on/ Staten Island./	Our camp/ is on/ Staten Island./
S2	Almost no campers/ have [ever] had/ a sunburn/ or caught a cold./	Almost every camper/ has [ever] had/ a sunburn/ or caught a cold./
S3	Since/ almost no campers/ have caught a cold,/ the parents are happy,/ and they praise the counselors./	Since/ almost every camper/ has caught a cold,/ the parents are unhappy,/ and they blame the counselors./
S4	They will/ use the camp/ again next year./	They will/ use the camp/ again next year./
S5	Is the camp in Staten Island?	Is the camp in Staten Island?