Natural Language Semantics manuscript No.
(will be inserted by the editor)

Non-monotonicity in NPI licensing

Luka Crnič

Received: date / Accepted: date

Abstract The distribution of the focus particle *even* is constrained: if it is adjoined at surface structure to an expression that is entailed by its focus alternatives, as in *even once*, it must be appropriately embedded to be acceptable. This paper focuses on the context-dependent distribution of such occurrences of *even* in the scope of non-monotone quantifiers. We show that it is explained on the assumption that *even* can move at LF (e.g., Karttunen and Peters 1979). The analysis is subsequently extended to occurrences of negative polarity items (NPIs) in these environments, which mirror the abovementioned distribution of *even* and which invalidate standard characterizations of NPI licensing conditions in terms of downward-entailingness. The idea behind the extension is that NPIs denote weak elements that are associates of covert *even* (e.g., Lee and Horn 1994). The paper concludes by discussing two comprehensive theories of NPI licensing and how our proposal relates to them.

Keywords *Even* · Non-monotonicity · Negative Polarity Items

1 Introduction

Many expressions in natural language have an idiosyncratic distribution. A prominent example of such an expression is the focus particle *even*. Its distribution is conditioned by whether the constituent in its immediate surface scope is entailed by its focus alternatives:[1] if it is entailed by its focus alterna-

Luka Crnič
Language, Logic and Cognition Center
The Hebrew University of Jerusalem
91905 Jerusalem, Israel
E-mail: luka.crnic@mail.huji.ac.il

[1] Entailment between objects of conjoinable type is defined in (i) (von Fintel 1999).

(i) *Cross-categorial entailment* ($\subseteq$)
a. For p, q of type t: p $\subseteq$ q iff p is false or q is true.

tives, the sentence is infelicitous unless *even* is appropriately embedded (e.g., Lahiri 1998). For example, if unembedded *even* associates with the focused element *once*, which is entailed by its alternatives *twice, thrice,* etc. (if something occurs twice, it occurs once, etc.), the sentence containing it is infelicitous and we say that the occurrence of *even* is unacceptable:

(1) #John read the book even ONCE.

In contrast, if an occurrence of *even* that associates with a weak element like *once* in its immediate surface scope – weak *even*, for short – is embedded under negation the sentence containing it is felicitous and we say that the occurrence of *even* is acceptable:

(2) John did not read the book even ONCE.

The above pattern parallels the well-known distribution of NPIs,[2] which are unacceptable if they are unembedded but acceptable under negation, as shown in (3) – a parallelism that has received considerable attention in the literature (see e.g. Heim 1984, Rooth 1985, Lee and Horn 1994, Krifka 1995, Lahiri 1998).

(3) a. #John read any book.
 b. John did not read any book.

More generally, if weak *even* and NPIs are in a downward-entailing (DE) environment, they are acceptable. We say that an element is in a DE environment if it is contained in a constituent that is DE with respect to them (Gajewski 2005, 2011, Homer 2012), as defined in the following:

(4) *Downward-entailing function:* A function f of type $\langle\delta,\tau\rangle$ is DE iff for all x, y of type δ such that $x \subseteq y$: $f(y) \subseteq f(x)$.

(5) *Downward-entailing environment:* A constituent **A** is DE with respect to α of type δ iff $\lambda x.\ [\![\mathbf{A}[\alpha/\mathbf{v}_\delta]]\!]^{g[v_\delta \to x]}$ is DE.

For illustration, if an expression is in the immediate scope of negation, it is in a DE environment; that is to say, if the expression is replaced with a stronger one, the truth of the sentence is preserved. More formally, the function that we obtain by abstracting over the expression is DE since the output of

b. For f, g of type $\langle\sigma,\tau\rangle$: $f \subseteq g$ iff for all x of type σ: $f(x) \subseteq g(x)$.

[2] NPIs may differ from each other with respect to their distribution. For example, so-called strong NPIs like *in weeks*, punctual *until*, and *either* can only occur in a subset of environments in which so-called weak NPIs like *any* and *ever* can occcur (see Gajewski 2011 for a recent discussion). We focus solely on the distribution of weak NPIs in this paper and plan to provide a more comprehensive account of the variation among NPIs elsewhere. For brevity, we will use the term 'NPIs' in the main text to refer to weak NPIs unless stated otherwise.

its application to a weaker element entails the output of its application to a stronger element. This is represented in (6), where abstraction over *once* yields a DE function.

(6) a. **[not [John read the book once]]**
b. λQ. ⟦**not [[John read the book] $\mathbf{t_1}$]**⟧$^{g[1\to Q]}$
c. that John did not read the book once ⊆ that John did not read the book twice, that John did not read the book thrice, etc.

1.1 Weak *even* in non-monotone environments

Another environment in which weak *even* may be acceptable is in the scope of a non-monotone quantifier like *exactly n NP*, i.e., in a non-monotone environment. However, weak *even* is not always acceptable in this environment. This is exemplified by the contrast between the sentences in (7) and (8), in which *even* associates with *once* and *opened*.[3]

(7) a. Exactly two congressmen read the constitution even ONCE.
b. Exactly four people in the whole world even OPENED that dissertation.

(8) a. #Exactly four hundred congressmen read the constitution even ONCE.
b. #Exactly one hundred people even OPENED that dissertation.

In contrast to the sentences in (8), sentences which have comparable assertive imports but in which weak *even* is in a DE environment are felicitous (note that there are 435 congressmen in the U.S. Congress):

(9) a. Exactly 35 congressmen didn't read the constitution even ONCE.
b. Many but not all people did not even OPEN that dissertation.

A characterization of the non-monotone environments in which weak *even* is acceptable is impossible without recourse to context – say, solely by reference to the make-up of the non-monotone quantifier. For example, a variant of the felicitous example in (7a), which features a different main predicate, may be infelicitous, even though weak *even* again occurs in the scope of a non-monotone quantifier with a low numeral:

(10) #Exactly two congressmen killed even ONE person.

[3] Note that *opened* is contextually but not logically entailed by its alternatives, *read* and *understood*; i.e., in all contexts compatible with our shared assumptions it holds that anyone who reads or understands a dissertation first opens that dissertation, even though this is not a logical necessity. We thus extend the notion of weakness to cover not just elements that are logically entailed by their alternatives but also elements that are contextually entailed by their alternatives.

Again, the counterpart of the sentence in (10) that has the same assertive import but in which weak *even* is in a DE environment is felicitous:

(11) Exactly 433 congressmen did not kill even ONE person.

Given these data, it is clear that weak *even* exhibits an intricate context-dependent behavior in non-monotone environments but not in DE environments – specifically, not in the scope of negation. Capturing the elusive context dependence of weak *even* in non-monotone environments presents the first challenge tackled in this paper:

(I) *Weak even under non-monotone quantifiers:* Provide an account of the distribution of weak *even* under non-monotone quantifiers.

1.2 NPIs in non-monotone environments

NPIs like *any* and *ever* may also occur in the scope of non-monotone quantifiers, as first observed by Linebarger (1980, 1987). Strikingly, their distribution in these environments parallels that of weak *even*; i.e., it appears to be sensitive to similar features of the context as that of weak *even*:

(12) a. Exactly four people in the whole world have ever read that dissertation: Bill, Mary, Tom, and Ed. (Linebarger 1987:373)
b. Exactly three students said anything in my seminar. (Gajewski 2008:73)

(13) a. #Exactly one hundred people have ever read that dissertation.
b. #Exactly ten of my twelve students said anything in my seminar.

The counterparts of the sentences in (13) that convey the same meaning but in which NPIs are in a DE environment are not context-dependent. This is exemplified in (14), which is a paraphrase of (13b).

(14) Exactly two of my twelve students did not say anything.

Similar to the distribution of weak *even*, a characterization of the non-monotone environments in which NPIs are acceptable is impossible without recourse to context. For example, a variant of the felicitous examples above in which weak *even* occurs in the scope of a non-monotone quantifier containing a low numeral, but now with a different predicate, may be infelicitous.

(15) #Exactly two congressmen killed anyone.

Again, the counterpart of the sentence in (15) that conveys the same meaning but in which weak *even* is in a DE environment is felicitous.

(16) Exactly 433 congressmen didn't kill anyone.

The felicitous occurrences of NPIs in non-monotone environments are particularly intriguing since they contradict the familiar assumption about NPI licensing (cf. Gajewski 2011, Homer 2012):[4]

(17) *NPI Licensing Condition*
An NPI is acceptable only if it occurs in a DE environment.

The second challenge, then, that we take on in this paper is to explain the context-dependent distribution of NPIs in the scope of non-monotone quantifiers. Due to their parallelism, a uniform explanation of the distributions of weak *even* and NPIs suggests itself and will be pursued.[5]

(II) *NPIs under non-monotone quantifiers:* Provide an account of the distribution of NPIs under non-monotone quantifiers.

1.3 A general theory of NPI licensing

A long-standing goal of linguistics is to provide a comprehensive, predictive, and compositional theory of the distribution of NPIs. For our analysis of NPIs in non-monotone environments to be considered adequate, we need to show that it is embeddable in such a general theory. One strategy that could be

[4] Although it is well known that the characterization of the NPI licensing condition in (17) is empirically inadequate in that it precludes NPIs from certain environments in which they are acceptable, say, the scope of *surprise* (see von Fintel 1999 for discussion), it is sufficient for the purposes of our paper, considering that the data examined here more generally undermine the idea that NPI licensing conditions can be characterized in a satisfactory manner (though see the discussion in Sect. 4.3).

[5] Although our focus in the main text is on non-monotone quantifiers of the form *exactly n NP*, there are other non-monotone quantifiers in whose scope one might expect to find weak *even* and NPIs. These include quantifiers of the form *between m and n NP* and *an odd/even number of NP*. The behavior of weak *even* and NPIs appears to be parallel in the scope of these expressions as well, though more speakers find them to be marked in these environments. For illustration, there is variation among speakers with respect to the sentences in (i)-(ii): while some speakers find all these sentences to be equally marked, others observe a clear contrast between them – the sentences in (i) are more acceptable than the sentences in (ii) (% represents variation in acceptability across speakers). On the other hand, all speakers find the sentences in (iii) unacceptable.

(i) a. %Between two and five congressmen read even ONE book last year.
b. %Between two and five congressmen read any book last year.

(ii) a. #Between three and four hundred congressmen read even ONE book last year.
b. #Between three and four hundred congressmen read any book last year.

(iii) a. #An odd number of congressmen read even ONE book last year.
b. #An odd number of congressmen read any book last year.

The account of the distribution of weak *even* and NPIs that we put forward in this paper provides us with the tools for explaining why there might be differences between non-monotone quantifiers with respect to how acceptable weak *even* and NPIs are in their scope and why there might be differences between speakers (see esp. footnote 10). (We discuss occurrences of weak *even* and NPIs in another potentially non-monotone environment in footnote 11.)

pursued in light of this task is to assume that the analysis, i.e., a uniform treatment of weak *even* and NPIs in non-monotone environments, extends to NPIs more generally, as stated in (18) (cf. Lee and Horn 1994, Lahiri 1998). This strategy is satisfactory to the extent that the common arguments against the descriptive adequacy of such a theory can be dispelled (cf. Heim 1984, Lahiri 1998).

(18) *Single operator theory of NPI licensing*
NPIs are weak elements that are associates of *even*.

Another strategy that one could pursue is to embed the analysis into a more elaborate theory of NPI licensing – say, a theory that takes NPIs to denote weak elements that are associates of either covert *even* or a covert exhaustification operator *exh*, as proposed in (20) (cf. Krifka 1995, Chierchia 2013).

(19) *Multiple operator theory of NPI licensing*
NPIs are weak elements that are associates of either *even* or *exh*.

The challenge, thus, can be summarized as follows:

(III) *Compatibility with a general theory of NPIs:* Show that the account of the distribution of NPIs in non-monotone environments is compatible with a descriptively and explanatorily adequate theory of NPIs.

We will first show that the single operator theory of NPI licensing is more viable than commonly assumed. Subsequently, we will show that our analysis can also be absorbed into the multiple operator theory, requiring only minor modifications of that theory. We remain neutral among the two theories.

1.4 Overview of the paper

The overarching goal of the paper is to show that the distribution of weak *even* and NPIs can be be explained by relying on the syntax and semantics of the expressions involved and their interaction with the context. More to the point, the paper takes on the three challenges described above and summarized in (20): we show that the context-dependent distribution of weak *even* in non-monotone environments can be explained on the movement approach to *even*; we argue that the distribution of NPIs in these environments can be explained analogously by assuming that it is governed by covert *even*; and we show how this assumption relates to two comprehensive theories of NPI licensing.

(20) *Summary of the challenges*
(I) Distribution of weak *even* under non-monotone quantifiers
(II) Distribution of NPIs under non-monotone quantifiers
(III) Compatibility with a general theory of NPI licensing

The paper has the following structure: Section 2 introduces some general assumptions about comparison of propositions and the scalar particle *even*, which crucially relies on such comparison. Following Lahiri's (1998) work on the distribution of weak *even* in DE environments, an explanation of the distribution of weak *even* in non-monotone environments is provided in Section 3. Section 4 shows that if we treat NPIs that occur in non-monotone environments as akin to weak *even* expressions, the explanation provided for the distribution of weak *even* naturally transfers to NPIs. Section 5 relates our proposal about the distribution of NPIs in non-monotone environments to two comprehensive theories of NPI licensing. We first explore the viability of a single operator theory that takes all NPIs to be associates of covert *even*, which is a trivial extension of our proposal (cf. Lee and Horn 1994, Lahiri 1998). We dispel some common arguments against this theory, arguments that are based on distinct distributions of weak *even* and NPIs in certain environments. Finally, we show that our proposal can also be naturally embedded in a multiple operator theory that takes NPIs to be associates of either covert *even* or a covert exhaustification operator (Krifka 1995, Chierchia 2013). Section 6 concludes the paper by presenting three avenues for future research.

2 Scalarity

A fundamental ability of human cognition is to make comparisons between objects along certain dimensions. Among the objects that can be compared are propositions. Comparisons of objects in general and propositions in particular are constrained by certain principles, which may be reflected in the distribution of operators sensitive to such comparisons. The following section describes these constraints and their reflexes as they pertain to the scalar particle *even*. The insights garnered in this section are used in subsequent sections to tackle the three challenges in (20).

2.1 Comparing propositions

Declarative sentences denote propositions (e.g., Stalnaker 1970). Propositions can be qualitatively compared, for instance, with respect to their informativity, remarkableness or likelihood. Consider the propositions in (21):

(21) a. John won the bronze medal.
b. John won the silver medal.
c. John won the gold medal.

These propositions may stand in different informativity, remarkableness, or likelihood relations to each other, which are conditioned on our assumptions about John and about winning medals. An example of such a relation is that John winning the gold medal is at most as likely as John winning the silver medal, which in turn is at most as likely as John winning the bronze medal, (22). (The notation '$p \leq_{\text{likely}} q$' or '$p \leq q$' stands for p being at most as likely as q; in the following, we only talk about likelihood for succinctness, though all our considerations transfer to informativity, remarkableness, etc.)

(22) a. that John won the gold medal
b. $\leq$ that John won the silver medal
c. $\leq$ that John won the bronze medal

If you hear someone say that John winning the gold medal is less likely than John winning the bronze medal, thereby expressing a relation compatible with (22), you may either object or modify your assumptions about John. But nothing about the nature of likelihood precludes such a relation from obtaining in at least some contexts. However, if someone goes on to say that John winning the bronze or silver medal is less likely than John winning the silver or gold medal, you would do well to point out the incoherence of her utterance.

(23) *Coherent comparison of propositions:* that John won the silver or gold medal $\leq$ that John won the bronze or silver medal $\Leftrightarrow$ that John won the gold medal $\leq$ that John won the bronze medal

The condition in (23) is a consequence of the objective constraint on the likelihood relation given in (24). It describes how likelihoods of mutually exclusive propositions condition likelihoods of disjunctions of such propositions.

(24) *Principle of coherence*
If propositions p, q, r are mutually incompatible, then the disjunction of p and q will be less likely than the disjunction of q and r iff p is less likely than r.

This constraint has important repercussions for any comparison of propositions that stand in an entailment relation. For example, consider the sentences in (25):

(25) a. John read the book once.
b. John read the book twice.
c. John read the book thrice.

These sentences stand in the entailment relation described in (26): that John read the book once is entailed by the propositions that John read the book twice and that John read the book thrice; it is the logically weakest of the three propositions.

(26) a. that John read the book once
b. $\supseteq$ John read the book twice
c. $\supseteq$ John read the book thrice

The principle of coherence in (24) dictates that a logically weaker proposition cannot be less likely than the logically stronger propositions. For illustration, that John read the book twice is, trivially, a disjunction of the proposition that John read the book twice and a contradiction, while that John read the book once is a disjunction of the propositions that John read the book twice and that John read the book exactly once. Since John reading the book exactly once is at least as likely as a contradiction, it follows from the principle of coherence that John reading the book once is at least as likely as John reading the book twice.

(27) a. that John read the book thrice
b. $\leq$ that John read the book twice
c. $\leq$ that John read the book once

More generally, the principle of coherence has the consequence that the logically stronger propositions are at most as likely as the logically weaker ones:[6]

(28) *Principle of entailment*
If a proposition p entails a proposition q, p is at most as likely as q.

Finally, in many conversational states various likelihood relations are compatible with the shared assumptions of the conversational participants. For example, as hinted at above, our shared assumptions may leave open, say, whether John winning a bronze medal is less likely than John winning a silver medal or whether it is equally likely. This is indicative of the fact that there may not be a unique likelihood function that is compatible with our shared assumptions. To capture this, we tentatively assume that common ground has more structure than is traditionally assumed: instead of treating it as a set of presuppositions in a context or, derivatively, as a set of worlds compatible with these presuppositions (context set), we may treat it as a set of pairs where the first element of a pair is a set of possible worlds (a live candidate for the context set) and the second element is a likelihood function defined on a Boolean algebra on that set of possible worlds (see blunt information states in Yalcin 2012). Accordingly, imposing a condition on the likelihood function/relation in conversation, say, by expressing a presupposition that John winning the bronze medal is less likely than John winning the silver medal, shifts the common ground to one in which all the pairs are such that their second element, the likelihood function, assigns a lower value to John winning the bronze medal than to John winning the silver medal; if there are no such pairs in the common ground, they are either accommodated or the conversation grounds to a halt. The common ground that we obtain may still contain many pairs whose second elements differ significantly from each other; for instance, according to some of them it may be less likely that John won the silver medal than that he won the gold medal, and according to some others the opposite may be the case. In the following, we will for simplicity assume that in a given conversational state a unique likelihood function/relation is provided, on which

[6] Many different kinds of measurements of propositions have been defended as appropriate representations of uncertainty and significance in formal epistemology. Many of them are more general than the probability or likelihood function that we rely on in this paper for concreteness. However, they all share the property in (28): they respect entailments. For a thorough discussion of these issues, see Halpern (2003), Chap. 2.

requirements may be imposed. This simplification should be understood as a shorthand for the common ground having to contain only pairs whose second element satisfies the respective requirement (see Yalcin 2011, 2012 for a more detailed discussion of the perspective on common ground sketched here).

2.2 Scalar particle *even*

Some expressions in natural language have meanings that rely on the qualitative comparison of salient propositions. An example of such an expression is the focus particle *even*, which we treat as a propositional operator that adjoins at the clause level. *Even* requires its propositional argument, its prejacent, to be less likely than all the relevant focus alternatives to the constituent it is adjoined to (Karttunen and Peters 1979, Wilkinson 1996, and many others); the likelihood relation is provided by the context. The assertive import of *even* is vacuous.[7]

(29) $[\![\textbf{even}_\textbf{C}]\!]^c$(p, w) is defined only if for all q in C: p ≠ q → p < q.
If defined, $[\![\textbf{even}_\textbf{C}]\!]^c$(p, w) = 1 iff p(w) = 1.

On this characterization of *even*, the sentence in (30) has a structure in which *even* associates with the focused phrase [*Syntactic Structures*]$_F$. The meaning of the sentence is defined in a context if John reading *Syntactic Structures* is less likely in that context than all the relevant alternatives, say, John reading *Alice in Wonderland*.

(30) a. John read even SYNTACTIC STRUCTURES.
b. **[even$_\textbf{C}$ [John read [Syntactic Structures]$_\textbf{F}$]]**
c. $[\![$(30b)$]\!]^c$ is defined only if for all relevant q in {that John read x: x is a book}: that John read Syntactic Structures < q, i.e., only if for all relevant books x other than Syntactic Structures: that John read Syntactic Structures < that John read x. If defined, $[\![$(30b)$]\!]^c$(w) = 1 iff John read Syntactic Structures in w.

The scalar presupposition in (30) is contingent: it is satisfied only in certain contexts and does not follow from axioms of probability theory or logic. If the presupposition triggered by *even* is not satisfied and if it cannot be easily accommodated, the sentence will be pragmatically marked.

An example of a sentence whose scalar presupposition is unsatisfiable is given in (31), where *even* associates with the weak element *once*; i.e., the

[7] There exists a weaker characterization of the scalar presupposition triggered by *even* than the one provided in (29) – one that requires the prejacent of *even* to be less likely than some of the relevant alternatives (e.g., Kay 1990, Bennett 1982); the arguments in the main text go through also on this characterization. We rely on the universal characterization primarily because of its prominence in the literature. Furthermore, another inference is sometimes taken to accompany *even*: the additive presupposition that there is an alternative to the constituent *even* is adjoined to that is true (see Guerzoni 2003, Sect. 2.7.3, for an overview). We defer the discussion of issues pertaining to additivity to another occasion.

sentence exemplifies an unembedded occurrence of weak *even*. The scalar presupposition of the sentence is that John reading the book once is less likely than all the relevant alternatives, say, John reading the book twice.

(31) a. #John read the book even ONCE.
b. **[even$_{\mathbf{C}}$ [John read the book once$_{\mathbf{F}}$]]**
c. $[\![(31b)]\!]^c$ is defined only if for all relevant n>1: that John read the book once < that John read the book n-times. (unsatisfiable)

This scalar presupposition is unsatisfiable because all the alternatives in the domain of *even* entail the prejacent of *even*, and so all of them are at most as likely as the prejacent, as stated in (32) – a consequence of the principle of entailment in (28). Accordingly, the sentence in (31) is infelicitous.

(32) a. *Entailment relation:* For all n≥1, that John read the book n-times ⊆ that John read the book once.
b. *Constraint on likelihood relations:* For all n≥1, that John read the book n-times ≤ that John read the book once.

As observed in the Introduction, the acceptability of weak *even* changes if it is embedded under negation. This is shown in (33a), which is a minimal modification of (31). All else being equal, this is unexpected on the proposal developed above: the prejacent of *even* continues to be entailed by the relevant alternatives in this environment and thus cannot be less likely than any of them. Since negation is a hole for presupposition projection, the deviant presupposition triggered by *even* should be inherited by the sentence. Seeing that the sentence in (33a) is felicitous, we are faced with a puzzle.

(33) a. John didn't read the book even ONCE.
b. **[not [even$_{\mathbf{C}}$ John read the book once$_{\mathbf{F}}$]]**
c. $[\![(33b)]\!]^c$ is defined only if for all relevant n>1: that John read the book once < that John read the book n-times. (unsatisfiable)

2.3 Movement of *even*

If *even* moves above negation at LF in the preceding example, we obtain the structure in (34) (see e.g. Karttunen and Peters 1979, Wilkinson 1996, Lahiri 1998 on the movement of *even*).

(34) a. John didn't read the book even ONCE.
b. **[even$_{\mathbf{C}}$ [not ~~even$_{\mathbf{C}}$~~ [John read the book once$_{\mathbf{F}}$]]]**

A consequence of this construal is that the alternatives in the domain of *even* now stand in the entailment relation described in (35): the prejacent of *even* is the logically strongest alternative, unlike in (33). All legitimate likelihood relations satisfy the constraint in (36).

(35) a. that John did not read the book once

b. $\subseteq$ that John did not read the book twice
c. $\subseteq$ that John did not read the book thrice, etc.

(36) *Constraint on likelihood relations:* For all n$\geq$1, that John did not read the book once $\leq$ that John did not read the book n-times.

The scalar presupposition triggered by *even* in (34) is compatible with the constraint in (36) and is thus consistent. Moreover, it is easily satisfied.

(37) $[\![$(34b)$]\!]^c$ is defined only if for all relevant n>1: that John did not read the book once < that John did not read the book n-times.

The presupposition is easily satisfied because in any context in which it is possible that John read the book exactly once, the propositions that John did not read the book once and that John did not read the book twice, etc. will have distinct likelihoods; the logically strongest proposition, the prejacent of *even*, will be the least likely (this is assuming that possibility in a context implies a non-zero degree of likelihood in that context). The construal of *even* above negation in (34) thus yields a structure whose meaning may well be felicitous (see Lahiri 1998 for discussion of the distribution of weak *even* in a variety of other DE environments). This approach to the distribution of *even* is suggestively called the 'movement approach':[8]

(38) *Movement approach to 'even'*
Even can move at LF and leave no trace.

Although the prejacent of *even* entailing all the relevant alternatives in the domain of *even* may suffice for *even* to be felicitous, as in (34), it is not

[8] The assumption that *even* can move at LF has received a lot of attention in the literature. We cannot do it justice here and refer the reader to a recent discusion in Nakanishi (2012) and the references cited therein. Nonetheless, it is worth pointing out that the assumption that movement of *even* leaves no trace, as stated in (38) and relied on in our representations, is endemic to our treatment of *even* as a propositional operator (cf. Rooth 1992). If instead we opted for the movement approach to focus association on which the focused phrase moves to the complement position of *even* (cf. Chomsky 1976, Drubig 1994, Wagner 2006), the movement of the *even*-phrase would leave a trace. On this approach, the sentence in (34) would have the structure in (i): *once* moves to the complement position of *even*, which subsequently moves above negation. We assume that the modifier *once* denotes a function from a set of events to a proposition that the set contains at least one event.

(i) $[\mathbf{even_C\ once_F}]_{\mathbf{(vt)(st)}}\ [_{\mathbf{(vt)}}\ \lambda \mathbf{y}\ [\mathbf{not}\ [\mathbf{y_v}\ \lambda \mathbf{x}\ \mathbf{John\ read\ the\ book\ x_v}]]]$

The meaning of *even* in (i) is the one given in (ii), where the types of its arguments are left unspecified; e.g., in (i) the x argument is a quantifier over events denoted by *once*.

(ii) $[\![\mathbf{even_C}]\!]^c$(Q, P, w) is defined only if for all alternatives Q' in C: Q'(P) $\neq$ Q(P) $\rightarrow$ Q(P) < Q'(P). If defined, $[\![\mathbf{even_C}]\!]^c$(Q, P, w) = 1 iff Q(P, w) = 1.

Since the meaning of the structure in (i) is indistinguishable from the one computed in the main text, (37), and since the representation, (34b), in the main text is simpler, we rely on it in the remainder of the paper for perspicuity.

a necessary condition for its felicity (see e.g. the felicitous example in (30) where the alternatives in the domain of *even* are logically independent). The necessary condition for the felicity of *even* is rather that no relevant alternative to the sister of *even* entails the prejacent of *even*. This condition has as a consequence the prediction that weak *even* may be acceptable if it occurs in the surface scope of a non-upward-entailing (non-UE) operator across which *even* can move at LF (viz. a non-monotone or a DE operator).

(39) *Prediction of the movement approach to 'even'*
Weak *even* may be acceptable if it occurs in the surface scope of a non-UE operator (a non-monotone or a DE operator).

Let us briefly elaborate on this prediction. On the one hand, if weak *even* occurs under a DE operator and *even* moves above it at LF, the prejacent of *even* will entail all the alternatives in its domain due to the entailment-reversing nature of DE operators. It will thus be at most as likely as other alternatives in the domain of *even* and in many natural contexts it will be less likely than the relevant alternatives. This has received extensive attention in the literature (e.g., Heim 1984, Lahiri 1998, Schwarz 2000). On the other hand, if weak *even* occurs under a non-monotone operator and *even* moves above it at LF, the prejacent of *even* will be logically independent of all the alternatives in its domain. It will be less likely than the relevant alternatives in some contexts. The main goal of the following section is to characterize the contexts in which this is the case for occurrences of weak *even* in the scope of non-monotone quantifiers, i.e., to tackle the first challenge presented in the Introduction, repeated below. Subsequently, we transpose our analysis to NPIs to account for their distribution under non-monotone quantifiers, which matches that of weak *even*.

(20) *Summary of the challenges*
(I) Distribution of weak *even* under non-monotone quantifiers
(II) Distribution of NPIs under non-monotone quantifiers
(III) Compatibility with a general theory of NPI licensing

3 Weak *even* under non-monotone quantifiers

The prediction of the movement approach to *even* is that weak *even* may be acceptable in the scope of non-monotone quantifiers: if *even* associates with a weak element in its immediate surface scope but moves across a non-monotone quantifier at LF, the focus alternatives to its sister in the target position are logically independent. Accordingly, it is conceivable that all of them that are relevant are more likely than its prejacent and that the scalar presupposition of *even* is satisfied, in which case the occurrence of *even* should be acceptable. This prediction is borne out. For example, the sentences in (40), in which weak *even* is in the surface scope of a non-monotone quantifier, are felicitous.

(40) a. Exactly two congressmen read the constitution even ONCE.
b. Exactly four people in the whole world even OPENED that dissertation.

According to the movement approach to *even*, these sentences can have the LFs in (41) where a non-UE operator intervenes between *even* and its weak associates *once* and *opened*, respectively:

(41) a. **[even$_{\mathbf{C}}$ [exactly two congressmen [read the const. once$_{\mathbf{F}}$]]]**
b. **[even$_{\mathbf{C}}$ [exactly four people [opened$_{\mathbf{F}}$ that dissertation]]]**

The alternatives in the domains of *even* in (41) are logically independent; in particular, none of the alternatives entails the prejacents of *even* except the prejacents themselves.

(42) a. that exactly two congressmen read the constitution once
b. $\not\supseteq$ that exactly two congressmen read the constitution twice, etc.

(43) a. that exactly four people opened that dissertation
b. $\not\supseteq$ that exactly four people read/understood that dissertation, etc.

And so the scalar presuppositions triggered by *even* in (41) are satisfiable, i.e., they comply with the principle of entailment. The scalar presupposition triggered by *even* in (41b) is provided in (44): it requires the prejacent of *even*, that exactly four people opened that dissertation, to be less likely than all the relevant alternatives, say, that exactly four people read that dissertation.

(44) $[\![$(41b)$]\!]^c$ is defined only if for all relevant x in {read, understand}: that exactly four people in the whole world opened that dissertation $<$ that exactly four people in the whole world x-ed that dissertation.

The presupposition in (44) is contingent. The first goal of this section is to explain why it and its ilk are satisfied in contexts compatible with our shared assumptions. Although we focus on the presupposition in (44), we outline the more general features of contexts in which scalar presuppositions triggered by weak *even* in the scope of non-monotone quantifiers of the form *exactly n NP* are satisfied. This goal is pursued in Sect. 3.1.

The felicity of weak *even* in the scope of non-monotone quantifiers changes if we tinker with (i) the numeral in the quantifier or (ii) the main predicate in the scope of the quantifier. The first form of tinkering is exemplified in (45): the only difference between these sentences and the sentences in (40) is in the numeral contained in the quantifier. (Although some speakers find the sentences in (40) slightly marked, they observe a clear contrast between them and the sentences in (45).)

(45) a. #Exactly four hundred congressmen read the constitution even ONCE.

b. #Exactly one hundred people even OPENED that dissertation.

On the basis of the contrast between (40) and (45), one could hypothesize that weak *even* in the surface scope of a non-monotone quantifier is acceptable if the numeral in the quantifier is low. However, such a hypothesis is refuted by examples like (46), which instantiate the second form of tinkering, where weak *even* is in the scope of the same quantifier as in (40a) but the sentence is nonetheless infelicitous.

(46) #Exactly two congressmen were caught taking a bribe even ONCE.

The second goal of this section, which parallels the first, is to explain why the presuppositions of the sentences in (45)-(46) are, though satisfiable, not satisfied in contexts compatible with our shared assumptions. This goal is pursued in Sect. 3.2. Accomplishing the two goals amounts to resolving the first challenge of the paper, i.e., providing an account of the distribution of weak *even* in the scope of non-monotone quantifiers.

3.1 Plausible presuppositions

We begin by explaining the plausibility of the presupposition in (44), that exactly four people in the whole world opening that dissertation is less likely than exactly four people in the whole world reading (or understanding, etc.) that dissertation. In doing this, we hope to point at the type of expectations that need to obtain for occurrences of weak *even* to be acceptable in the scope of non-monotone quantifiers of the form *exactly n NP* more generally.

Intuitively, the reason why the presupposition in (44) is satisfied in contexts compatible with our shared assumptions or is, at least, easily accommodable in such contexts lies in our expectations about people standing in a particular relation to that dissertation. Specifically, we conceivably share the expectations with respect to a dissertation in, say, theoretical linguistics: (i) that many people opened it, perhaps about fifty; (ii) that fewer people read it, perhaps about twenty; and (iii) that only few people understood it, perhaps about ten.

(47) *Approximation of our expectations*
- a. the number of people who opened the dissertation ≈ high (50)
- b. the number of people who read the dissertation ≈ moderate (20)
- c. the number of people who understood the dissertation ≈ low (10)

These expectations serve as a guidepost for what qualifies as a plausible likelihood ordering on the alternatives. That is to say, we abduce the relative likelihoods of the alternatives under consideration from these expectations and some innocuous assumptions about how they are distributed. Roughly, it is less likely that exactly four people opened that dissertation than that exactly four people read/understood that dissertation because (a) the disparity between exactly four people opening that dissertation, which is a relatively low number, and the expected fifty people doing it, which is a relatively high number, is considerably greater than (b) the disparity between exactly exactly four people reading/understanding that dissertation and the expected twenty/ten people doing it, which are relatively low numbers.

More to the point, if we expect that about fifty people opened that dissertation, then we take it to be less likely that many fewer than fifty people opened that dissertation, say exactly twenty, than that fifty people opened it. And we take it to be less likely that almost no people opened that dissertation, say, exactly four, than that exactly twenty people did etc. This means that our expectation about exactly n people opening that dissertation increases as n approaches fifty. Similar reasoning applies to our expectation about exactly n people reading that dissertation, which increases as n approaches twenty, and to our expectation about exactly n people understanding that dissertation, which increases as n approaches ten. To help our understanding, we graphically represent our expectations by example probability distributions in Figure 1: the modes of the distributions corresponding to opening, reading, and understanding that dissertation are at 50, 20, and 10, respectively, and we assume that the distributions are approximately normal (which is plausible, though not crucial for our reasoning).[9]

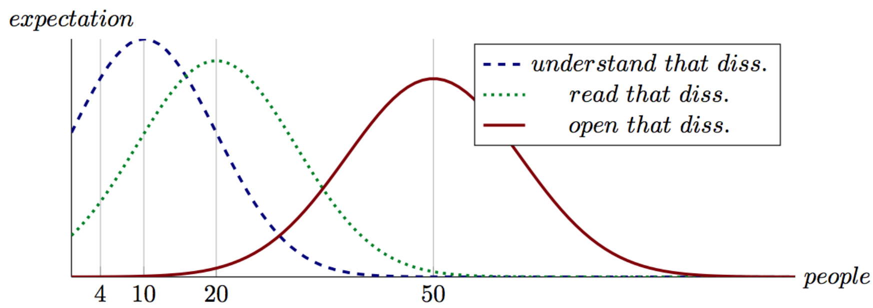

Fig. 1 An illustration of our expectations. The x-axis stands for the number of people, the y-axis for the measure of expectedness. Graphs correspond to the alternatives.

[9] Our assumptions pertaining to the representation of likelihood and how it figures in the evaluation of the scalar presupposition of *even* are perhaps oversimplistic. For example, a reviewer suggests that instead of probability functions, one might want to employ cumulative distribution functions in characterizing the scalar presupposition. They draw support for this suggestion from the apparent felicity of the sentence *Exactly four people in the whole world even OPENED that dissertation* in contexts in which many people are expected to open the disseration but in which it is known that only two people are qualified to understand the dissertation (or have the time and patience to read it), say, the author's advisors. In such a context the likelihood of the prejacent of *even* would be ex hypothesi at least as great as that of the alternatives built on *understood* (and *read*). Accordingly, the scalar presupposition of the sentence should not be satisfied and the sentence should be infelicitous. A switch to, say, cumulative distribution functions would avoid this issue because one would effectively compare the likelihood that exactly four or fewer people opened that dissertation with the likelihood that exactly four or fewer people read/understood it, and in the described context the former would be lower than the latter. Since a proper exploration of how precisely to model measurement of propositions that *even* relies upon would require more space than we can allot to it, we shelve it for another occasion and continue to rely on probability functions in our examples in the main text (see also footnote 6).

On the basis of the above reasoning, the likelihood relations that are compatible with our shared assumptions arguably all satisfy the following condition:

(48) *Plausible likelihood relations:* that exactly four people opened that dissertation $<$ that exactly four people read that dissertation $\leq$ that exactly four people understood that dissertation

Accordingly, it plausibly holds that exactly four people opening that dissertation is less likely than all the relevant alternatives. The scalar presupposition triggered by *even*, repeated below, is thus satisfied.

(44) $[\![(41b)]\!]^c$ is defined only if for all relevant x in {read, understand}: that exactly four people in the whole world opened that dissertation $<$ that exactly four people in the whole world x-ed that dissertation.

More generally, the scalar presupposition triggered by *even* that associates with a weak element in the scope of a non-monotone quantifier *exactly n NP* is satisfied only if the numeral in the non-monotone quantifier is (i) appropriately lower than the expected number of individuals that are both in the domain of the quantifier and in the denotation of the main predicate – e.g., the expected number of people opening that dissertation – and (ii) appropriately close to the the expected number of individuals that are in the domain of the non-monotone quantifier and in the relevant stronger alternatives to the main predicate – e.g., the expected number of people reading/understanding that dissertation. What counts as appropriately lower or appropriately close depends on the context, i.e., on the shared expectations of conversational participants. If the two conditions in (i) and (ii) do not obtain, the scalar presupposition will not be satisfied and either the sentence will be perceived as infelicitous or appropriate assumptions will be accommodated. We discuss violations of these conditions in the following.[10]

3.2 Implausible presuppositions

There are two types of examples that we brought up in the preceding discussion in which weak *even* is unacceptable in the scope of non-monotone quantifiers.

[10] In footnote 5 we observed that there appear to be differences between non-monotone quantifiers with respect to how acceptable weak *even* is in their scope. For example, some speakers find weak *even* to be more marked in the scope of *between m and n NP* than in the scope of *exactly n NP*, though most speakers do observe a clear contrast between the sentences in (i): the sentence in (ia) is more acceptable than the sentence in (ib).

(i) a. %Between two and five congressmen read even ONE book last year.
b. #Between three and four hundred congressmen read even ONE book last year.

We suggest that the source of this difference between, say, *exactly n NP* and *between m and n NP* lies in the difference in how difficult it is to extrapolate the required likelihood relations from our shared assumptions and the nature of the probability distributions involved.

The first type involves sentences like the one in (49a), where weak *even* is in the scope of a non-monotone quantifier that contains a relatively high number. The sentence may have the LF in (49b), where *even* associates with the weak element *opened* and has moved above the non-monotone quantifier at LF, a configuration that parallels the ones of the felicitous sentences in (40).

(49) a. #Exactly one hundred people even OPENED that dissertation.
b. **[even$_{\mathbf{C}}$ [exactly 100 people opened$_{\mathbf{F}}$ that dissertation]]**

The scalar presupposition of (49b) is consistent for the same reason that the presupposition in (44) is consistent: the domain of *even* contains logically independent alternatives.

(50) $[\![$(49b)$]\!]^c$ is defined only if for all relevant x in {read, understand}: that exactly one hundred people opened that dissertation $<$ that exactly one hundred people x-ed that dissertation.

However, if our expectations are those described in (47) – that about fifty people opened that dissertation, that about twenty people read it, and that about ten people understood it – then the proposition that exactly one hundred people opened that dissertation will be closer to our expectations than the proposition that exactly one hundred people read or understood it. That is, roughly, (a) the disparity between exactly one hundred people opening that dissertation, which is a relatively high number, and the expected fifty people opening it, which is also a relatively high number, is lower than or equals (b) the disparity between exactly one hundred people reading/understanding it and the expected twenty/ten people reading/understaning it. This can also be gleaned from Figure 1: the expectedness of exactly one hundred people reading or understanding that dissertation is lower than (or perhaps indistinguishable from) the expectedness of exactly one hundred people opening that dissertation. Accordingly, it cannot be the case that the latter proposition, which conforms to our expectations better than the alternatives, is less likely than the alternatives. That is, the plausible likelihood relations on the relevant alternatives satisfy the condition in (51). Since this condition clashes with the presupposition in (50), the presupposition is not satisfied in contexts compatible with our shared assumptions and the sentence in (49) is perceived to be infelicitous.

(51) *Plausible likelihood relations:* that exactly 100 people understood that dissertation $\leq$ that exactly 100 people read that dissertation $\leq$ that exactly 100 people opened that dissertation

The second type of example where weak *even* is unacceptable in the scope of non-monotone quantifiers involves main predicates that are expected to hold of only very few objects in the domain of the non-monotone quantifier and whose alternatives are expected to hold of even fewer such objects. An example of this is provided in (52), where the non-monotone quantifier is identical to the one in the felicitous example in (40a).

(52) a. #Exactly two congressmen were caught taking a bribe even ONCE.
b. **[even$_{\mathbf{C}}$ [exactly two congressmen were caught taking a bribe once$_{\mathbf{F}}$]]**

The scalar presupposition triggered by *even* in (52) is described in (53): that exactly two congressmen were caught taking a bribe once is less likely than all the relevant alternatives – say, that exactly two congressmen were caught taking a bribe twice.

(53) $[\![(52b)]\!]^c$ is defined only if for all relevant n>1: that exactly two congressmen were caught taking a bribe once < that exactly two congressmen were caught taking a bribe n-times.

Now, it generally holds that if a politician is caught taking a bribe, his career is effectively over. This means that it is practically impossible for a congressman to be caught taking a bribe more than once. In accordance, politicians take great precautions against being caught taking bribes. These precautions translate into an expectation that very few congressmen – perhaps two or three – are caught taking a bribe. Our expectations are given in (54) and graphically represented in Figure 2.

(54) *Approximation of our expectations*
a. the number of congressmen getting caught once ≈ 3
b. the number of congressmen getting caught twice ≈ 0
c. the number of congressmen getting caught thrice ≈ 0

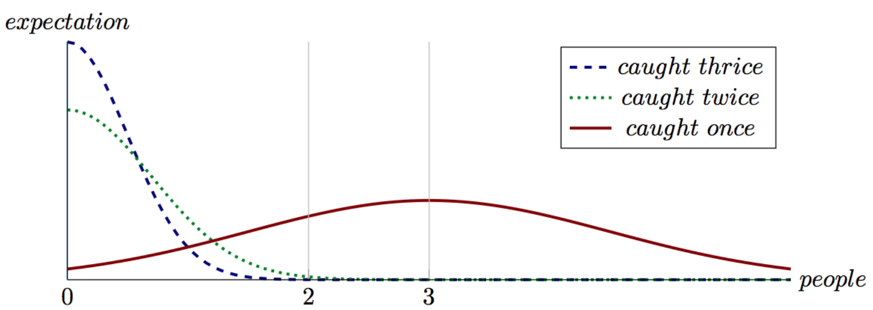

Fig. 2 An illustration of our expectations. The x-axis stands for the number of people, the y axis for a measure of expectedness. Graphs correspond to the alternatives.

Consequently, as stated in (55), it cannot be less likely that exactly two congressmen were caught taking a bribe once than that exactly two congressmen were caught taking a bribe twice (or more often), a practical impossibility. This means that the scalar presupposition in (53) is not satisfied in contexts compatible with our shared assumptions, explaining the infelicity of (52).

(55) *Plausible likelihood relations:* that exactly two congressmen were caught taking a bribe twice etc. $\leq$ that exactly two congressmen were caught taking a bribe once

Finally, sentences that are comparable to the ones discussed in this subsection in that they convey the same assertive meanings but that differ from them in that weak *even* is in a DE environment are felicitous. This too follows straightforwardly from the movement approach to *even*. For example, the sentence in (56a) has the same assertive import as the sentence in (52). Its structure is provided in (56b), where *even* associates with the weak element *once* across negation but crucially takes scope below the quantifier.

(56) a. Exactly 433 congressmen were not caught taking a bribe even ONCE.
b. **[exactly 433 congressmen]$_{\mathbf{x}}$ even$_{\mathbf{C}}$ [not [x were caught taking a bribe once$_{\mathbf{F}}$]]**

The scalar presupposition triggered by *even* in (56) is that the assignment-dependent proposition that x was not caught taking a bribe once is less likely than that x was not caught taking a bribe n-times for all relevant n>1. Assuming either existential or universal presupposition projection in the scope of non-monotone quantifiers, we obtain a scalar presupposition for the sentence, given in (57), that is easily satisfied in contexts compatible with our shared assumptions, since we take it to be practically certain that congressmen do not get caught taking a bribe twice, though they might get caught taking a bribe exactly once. (Similar considerations hold if the presupposition projected from the scope of a non-monotone quantifier has a different quantificational force from what is assumed in (57).)

(57) For some/every congressman x, for all relevant n>1, that x was not caught taking a bribe once $<$ that x was not caught taking a bribe n-times.

The acceptability of the sentence in (56) is thus correctly predicted to be unproblematic: *even* may take scope below the non-monotone quantifier and trigger a scalar presupposition that is satisfied in contexts compatible with our shared assumptions. Clearly, this is is not possible in sentences in which no DE operator is present in the structure.

To summarize: we have shown that the prediction of the movement approach to *even* is confirmed by the distribution of weak *even* in the scope of non-monotone quantifiers. Moreover, we have shown that the scalar semantics of *even* allows us to correctly predict the context dependence of weak *even* in the scope of non-monotone quantifiers: it is acceptable only if appropriate expectations obtain in the context. More specifically, we have indicated that it has to hold that the expected number of individuals that are in the main predicate of the prejacent and in the restrictor of the non-monotone quantifier must be (i) appropriately greater than the number used in the non-monotone quantifier and (ii) appropriately greater than the expected number of individuals that are in the relevant alternatives to the main predicate and in the restrictor of the non-monotone quantifier. If these two conditions are not met, the scalar presupposition of the respective sentence is not satisfied and the sentence is pragmatically marked. We have thus resolved the first of the three challenges that we intend to tackle in this paper, repeated below.

(20) *Summary of the challenges*
 (I) Distribution of weak *even* under non-monotone quantifiers ✓
 (II) Distribution of NPIs under non-monotone quantifiers
 (III) Compatibility with a general theory of NPI licensing

In the following section we turn to the second challenge. We extend the analysis of weak *even* under non-monotone quantifiers to NPIs like *any* and *ever*, building primarily on Lee and Horn's (1994) proposal for *any* and Lahiri's (1998) proposal for Hindi NPIs.

4 NPIs under non-monotone quantifiers

It has been observed that NPIs may felicitously occur in the scope or the restrictor of non-monotone quantifiers:

(58) a. Exactly four people in the whole world have ever read that dissertation: Bill, Mary, Tom, and Ed. (Linebarger 1987:373)
 b. Exactly three people with any money showed up. (Rothschild 2006:229)
 c. Exactly three students said anything in my seminar. (Gajewski 2008:73)

These examples invalidate extant characterizations of NPI licensing conditions in terms of downward-entailingness, such as stated in (17), repeated below. In these examples the scope of non-monotone quantifiers is not a DE environment. But they are usually assigned a peripheral role in the discussion of NPI licensing (Linebarger 1987 and Rothschild 2006 are notable exceptions).

(17) *NPI Licensing Condition*
 An NPI is acceptable only if it occurs in a DE environment.

The marginality of the discussion of NPIs in non-monotone environments is at least partly due to the elusiveness of their felicity conditions (but see Linebarger 1987). For example, a variant of the sentence in (58a) in which the numeral in the non-monotone quantifier is replaced with *one hundred* is marked. (Or else it leads us to accommodate the implausible assumption that a lot more than one hundred people are expected to have read that dissertation.)

(59) #Exactly one hundred people have ever read that dissertation.

The example in (60) illustrates that it is not only the size of the numeral by itself that matters for the felicity of the NPI but also the content of the main predicate hosting the NPI, suggesting that a context-independent characterization of NPI licensing conditions is indeed untenable.

(60) #Exactly two congressmen killed anyone last year.

In light of these data, we are faced with two tasks: an account of NPIs needs to be provided that (i) is flexible enough to allow them to occur in the scope of non-monotone quantifiers and (ii) captures their context dependence in these environments. We propose that the mechanism governing the distribution of NPIs in the scope of non-monotone quantifiers is the same mechanism that underlies the distribution of weak *even*; i.e., NPIs are weak existential quantifiers whose distribution in non-monotone environments is regulated by a covert *even* that associates with them (cf. Krifka 1995 on stressed NPIs; Schmerling 1971, Heim 1984 on minimizers; and Lahiri 1998 on Hindi NPIs).[11]

4.1 Some assumptions about NPI licensing

Explanatory approaches to NPI licensing aim at accounting for the distribution of NPIs in terms of their syntactic-semantic properties in combination with the properties of the environments in which they (cannot) occur. A groundbreaking discussion of NPIs by Kadmon and Landman (1993) inspired several such approaches that converged on three assumptions (see esp. Krifka 1995, Lahiri 1998, Chierchia 2013): (i) that NPIs denote existential quantifiers; (ii) that they induce alternatives, similarly to focused elements; and (iii) that the alternatives they induce are utilized by specific alternative-sensitive operators. We adopt these assumptions in our treatment of NPIs under non-monotone

[11] There is another class of poorly understood occurrences of NPIs that might be explained in a similar manner, namely, NPIs that occur in the scope of desire predicates and in imperatives. They are exemplified in (i); parallel examples with weak *even* are in (ii) (see Kadmon and Landman 1993 for discussion of NPIs under emotive factive predicates).

(i) a. Mary hopes to ever make a video of that quality.
b. I am glad that our representative read any book at all.
c. Show me a political party that ever cared for the people!

(ii) a. Mary hopes to make even ONE video of that quality.
b. I am glad that our representative read the constitution even ONCE.
c. Show me even ONE political party that cares for the people!

The felicitous occurrences of NPIs and weak *even* in (i)-(ii) are unexpected if these environments are UE (e.g., Hintikka 1962, von Fintel 1999, Schwager 2005). However, the data can be explained on non-monotone analyses of desire predicates and imperatives, e.g., the negation-related analyses of Heim (1992), Villalta (2008), or Lassiter (2011). If we assume these analyses, it is expected that weak *even* and NPIs (as weak associates of *even*) may be felicitous in desire statements and in imperatives. Moreover, the scalar presupposition triggered by *even* in these configurations could be capitalized on to explain the protean context dependence of NPIs in these environments (cf. Kadmon and Landman 1993). An exploration of these issues is left for future work (see Crnič 2013b for an overview).

quantifiers. That is, we assume that NPIs like *any* and *ever* induce alternatives that are stronger than the meaning of the NPI and that these alternatives can be used up by covert *even*, as proposed by Lee and Horn (1994) and by Lahiri (1998) for Hindi NPIs. We discuss the details of these assumptions in the following. (We relate our proposal to more comprehensive theories of NPI licensing in Sect. 5.)

Existential quantification

Our first assumption is that NPIs like *any* and *ever* are indefinites, an assumption that is uncontroversial (see Gajewski 2008 for an evaluation of the main arguments for and against it). We thus treat *any* as denoting an existential quantifier that takes a predicate denoted by its sister (restrictor) and a predicate denoted by the sister to the DP (nuclear scope) as its arguments and contributes the meaning that there exists an individual of which both the restrictor and the nuclear scope hold. (For reasons of notational simplicity, we ignore the resource domain argument of the quantifier until Sect 5.2.)

(61) $[\![\textbf{any}]\!]^c(\text{P, Q}) = [\![\textbf{one}]\!]^c(\text{P, Q}) = 1$ iff $\exists$x [P(x) = Q(x) = 1].

On this assumption, the truth-conditional contribution of *any* to a clause containing it, as analyzed in example (62) below, corresponds to the contribution of other indefinites. (In subsequent LFs, we represent NPIs in situ for readability.)

(62) a. #John read any book.
b. **[any book]$_\mathbf{x}$ [John read x]**
c. $[\![(62\text{b})]\!]^c(\text{w}) = 1$ iff $\exists$x [x is book & John read x in w].

The meaning computed in (62c) is licit. However, the sentence in (62a) that is supposed to have this meaning is unacceptable. This is a consequence of the alternatives induced by the NPI and of how these alternatives are used.

Alternatives

Our second assumption is that NPIs obligatorily induce alternatives that are stronger than the NPIs. We implement this assumption by taking the alternatives induced by NPIs to be number indefinites (Heim 1984 and Lahiri 1998 also entertain this assumption; see Sect. 5.2 for an alternative implementation).[12] For example, the alternatives to *any* are given in (63) and their pointwise composition with a minimal clause is in (64).

(63) ALT(**any**) = {one, two, three, ...}

[12] The choice of taking alternatives to NPIs to be number indefinites is not essential and is motivated primarily by the desire to make the parallelism between NPIs and weak *even* more transparent. Our proposal is compatible with other implementations (see Sect. 5.2).

(64) ALT(**John read any book**) = {that John read n books: n≥1}

It clearly holds that *any* is entailed by all of its alternatives and that minimal clauses containing *any* are entailed by the alternatives built from the alternatives to *any* – NPIs are elements that are weaker than their alternatives.

(65) For all p ∈ ALT(**John read any book**): p ⊆ that John read a book; i.e., for all n≥1: that John read n books ⊆ that John read a book.

Covert even

Our final assumption relates to how these alternatives are used up in the sentences under discussion, i.e., in sentences in which NPIs occur in the scope of non-monotone quantifiers. We assume that in these sentences NPIs are accompanied by covert *even* (see e.g. Lee and Horn 1994, Lahiri 1998 for the assumption that certain NPIs are accompanied by *even*).

(66) *Licensing of NPIs in non-monotone environments*
NPIs in non-monotone environments are associates of covert *even*.

The semantics of covert *even* is identical to that of overt *even* and is repeated below. (We do not employ special notation for covert *even* in our LF representations and rely on the reader to deduce its PF status.)

(29) $[\![\textbf{even}_\textbf{C}]\!]^c$(p, w) is defined only if for all q in C: p ≠ q → p < q.
If defined, $[\![\textbf{even}_\textbf{C}]\!]^c$(p, w) = 1 iff p(w) = 1.

Given these assumptions, a minimal unembedded clause containing an NPI would have a representation along the lines of (67), where *any* is in the immediate scope of covert *even*.

(67) a. #John read any book last year.
b. [**even**$_\textbf{C}$ [**John read any book last year**]]

The interpretation of the structure in (67) yields an inconsistent meaning – specifically, an inconsistent scalar presupposition. This is because the domain of *even* in (67) consists only of alternatives that entail the prejacent, (68), and the presupposition of *even* is that the prejacent is less likely than all the alternatives in the domain of even, as stated in (69). This is a requirement that is in conflict with the principle of entailment and thus unsatisfiable.

(68) C ⊆ {that John read n books last year: n≥1}

(69) $[\![$(67b)$]\!]^c$ is defined only if for all relevant n>1: that John read a book last year < that John read n books last year. (unsatisfiable)

However, if an NPI is appropriately embedded at surface structure, say, under negation, the sentence is felicitous. The reason for this is that, unlike in the case of (69), the sentence allows for a construal on which *even* takes scope above negation:

(70) a. John didn't read any book last year.
b. [**even$_C$** [**not** [**John read any book last year**]]]

The domain of *even* in (70) consists of the alternatives described in (71), all of which are entailed by the meaning of the prejacent in (70): if John did not read a book last year, then he did not read two or more books last year.

(71) C ⊆ {that John did not read n books last year: n≥1}

Accordingly, the scalar presupposition triggered in (70) is satisfiable – it complies with the principle of entailment. And it is satisfied in any context in which it is possible that John read exactly one book; in any such context the proposition that John did not read one book is not contextually equivalent to John not reading two books but rather asymmetrically contextually entails it and is thus less likely than it.

(72) $[\![$(70b)$]\!]^c$ is defined only if for all relevant n>1: that John did not read a book last year < that John did not read n books last year.

Sentences containing NPIs in the scope of a DE operator thus easily trigger licit scalar presuppositions. We see that the distribution of NPIs in DE environments can be cogently captured on the assumptions we made above (see Lahiri 1998 for further discussion). However, the goal of this section is to deal with NPIs in the scope of non-monotone quantifiers.

4.2 Non-monotone environments

The acceptability of NPIs in non-monotone environments is highly context-dependent, as illustrated by the contrast in felicity of the sentences in (73). We show in the following that this context dependence is explained by our assumptions about the licensing of NPIs in these environments. Predictably, the explanation mirrors our explanation of the distribution of weak *even*.

(73) a. Exactly two congressmen read any book last year.
b. #Exactly four hundred congressmen read any book last year.
c. #Exactly two congressmen killed anyone last year.

The sentence with an acceptable occurrence of *any* in (73a) has the construal in (74), where *even* is adjoined to the matrix clause while the NPI is interpreted in the scope of the non-monotone quantifier.

(74) a. Exactly two congressmen read any book last year.
b. [**even$_C$** [**exactly two congressmen read any book**]]

Since the NPI is in the scope of a non-monotone quantifier, it holds that the alternatives induced by it are logically independent. This means that the scalar presupposition triggered by *even* in (74) is satisfiable – that exactly two congressmen read a book last year may be less likely than, say, that exactly two congressmen read two books last year. The assertive import of the sentence on this construal is that exactly two congressmen read a book last year.

(75) $[\![(74b)]\!]^c$ is defined only if for all relevant n>1: that exactly two congressmen read a book last year < that exactly two congressmen read n books last year.

Plausible presuppositions

A consequence of the logical independence of the alternatives in the domain of *even* in (74) is that the scalar presupposition in (75) is contingent: it is satisfied only in certain contexts. Intuitively, it is satisfied in contexts that are compatible with our shared expectations. Namely, we expect that almost all congressmen read a book last year, say, about four hundred, and that many of these read at least five books, say, about two hundred, and that a moderate number of them read at least ten books, say, about fifty. These expectations are summarized in (76) and represented in Figure 3.

(76) *Approximation of our expectations*
  a. the number of congressmen reading a book ≈ very high (400)
  b. the number of congressmen reading five books ≈ high (200)
  c. the number of congressmen reading ten books ≈ moderate (50)

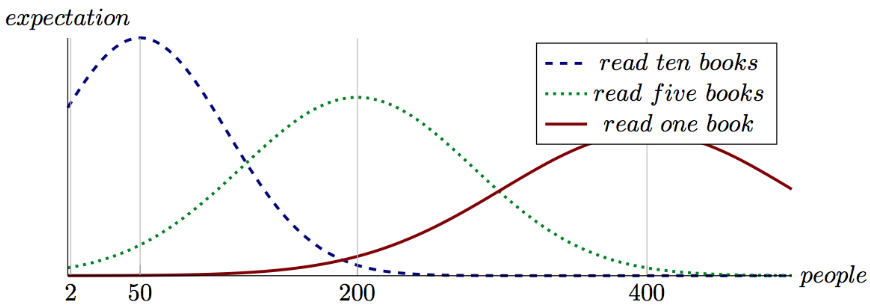

Fig. 3 An illustration of our expectations. The x-axis stands for the number of people, the y-axis for a measure of expectedness. Graphs correspond to the alternatives.

The prejacent of *even* in (74) is extremely unlikely in light of our expectation that almost all congressmen read at least one book. And it is less likely than that exactly two congressmen read five/ten books, which are alternatives that conform better to our expectations, namely, that about two hundred/fifty congressmen read five/ten books, respectively. This means that if the domain of *even* is restricted to these three alternatives, the scalar presupposition of the sentence in (75) is satisfied. We can say that the NPI is ‘licensed.’

Implausible presuppositions

If either the numeral in the non-monotone quantifier or the main predicate are modified, the sentence may become infelicitous due to a shift in pertinent expectations. For example, the scalar presupposition triggered by *even* in (77) is false in contexts where we expect, as described above, that about four hundred congressmen read a book last year.

(77) a. #Exactly four hundred congressmen read any book last year.
 b. **[even$_C$ [exactly 400 congressmen read any book]]**
 c. $[\![(77b)]\!]^c$ is defined only if for all relevant n>1: that exactly 400 congressmen read a book last year < that exactly 400 congressmen read n books last year.

It holds that since we expect about four hundred congressmen to read a book, we expect that at most four hundred congressmen read two books last year and that even fewer congressmen read three books, etc. In accordance with these expectations, it is quite likely that the prejacent of *even* is true, i.e., that exactly four hundred congressmen read a book last year, and it is at most as likely that some alternative is true, say, that exactly four hundred congressmen read two books. This means that the scalar presupposition in (77c) is not satisfied in contexts compatible with our shared assumptions and thus the NPI is not licensed.

Finally, the sentence in (78) triggers the scalar presupposition that the proposition that exactly two congressmen killed someone last year is less likely than all the relevant alternatives, say, that exactly two congressmen killed two people last year.

(78) a. #Exactly two congressmen killed anyone last year.
 b. **[even$_C$ [exactly two congressmen killed any person]]**
 c. $[\![(78b)]\!]^c$ is defined only if for all relevant n>1: that exactly two congressmen killed a person last year < that exactly two congressmen killed n people last year.

It arguably holds that we do not expect any congressmen to kill anyone. Accordingly, that exactly two congressmen killed two (or more) people cannot be more likely than that exactly two congressmen killed someone. This means that the prejacent of *even* in (78) is at least as likely as the alternatives and so the scalar presupposition is not satisfied, causing the NPI not to be licensed.[13]

[13] A reviewer reports a contrast in acceptability between the sentences in (i), which might appear to run afoul of the claimed parallelism between weak *even* and NPIs.

(i) a. #Exactly two congressmen were caught accepting a bribe even ONCE.
 b. Exactly two congressmen were ever caught accepting a bribe.

However, this difference in acceptability is expected on our proposal, at least if *ever* is not taken to mean *(even) once* (cf. Heim 1984) but rather, approximately, *(even) during the longest potentially relevant period of time* (cf. Krifka 1995). Indeed, while we may expect that only few congressmen were caught taking a bribe during some salient period of time – say, two

NPIs in sentences that convey the same meaning as those in (77)-(78) but in which the NPIs are in a DE environment are acceptable. This is predicted on the assumption that NPIs are associates of *even*. For example, the sentence in (79a) may have the structure in (79b), in which *even* takes scope above negation but below the non-monotone quantifier.

(79) a. Exactly thirty-five congressmen did not read any book last year.
b. **[exactly 35 congressmen]$_{\mathbf{x}}$ [even$_{\mathbf{C}}$ [not [x read any book]]]**

The scalar presupposition which *even* triggers in (79) is assignment-dependent: that x did not read a book is less likely than that x did not read n books for all relevant n>1. On any reasonable assumption about how this presupposition projects in the scope of non-monotone quantifiers, we obtain a global presupposition, (80), that is easily satisfied in contexts compatible with our shared assumptions. By contrast, since in (77) and (78) no scope site above a DE operator is available at which *even* could be interpreted, *even* must be interpreted above the non-monotone quantifier, where it gives rise to, as we have seen, a more involved scalar presupposition than (80).

(80) For some/every congressman x, for all relevant n>1, that x did not read a book < that x did not read n books.

To summarize: we have shown that the explanation that was provided for the distribution of weak *even* in Sect. 3 translates naturally to NPIs if we assume that NPIs in the scope of non-monotone quantifiers denote weak elements that are associates of covert *even*. In such configurations, *even* triggers a scalar presupposition that is satisfied only in contexts in which appropriate expectations obtain or can easily be accommodated. This resolves the second of the three challenges tackled in this paper, repeated below.

(20) *Summary of the challenges*
(I) Distribution of weak *even* under non-monotone quantifiers ✓
(II) Distribution of NPIs under non-monotone quantifiers ✓
(III) Compatibility with a general theory of NPI licensing

We conclude the section by studying a variant of existing NPI licensing conditions and checking whether it can adequately capture the observed context-dependent distribution of NPIs under non-monotone quantifiers.

or three last year – as described in (54), we may expect that a significantly greater number of congressmen were caught during some longer period of time – say, since Congress was established. Accordingly, it may well be less likely that exactly two congressmen were caught accepting a bribe during this long period of time than, say, that exactly two congressmen were caught accepting a bribe last year.

4.3 Modified NPI Licensing Condition

Linebarger (1987) notes in her discussion of NPIs in the scope of non-monotone quantifiers that they invalidate the characterization of the NPI Licensing Condition in terms of downward-entailingness:

(17) *NPI Licensing Condition*
An NPI is acceptable only if it occurs in a DE environment.

We could replace this condition with the weaker one in (81):

(81) *Modified NPI Licensing Condition* (version 1 of 2)
An NPI is acceptable only if it occurs in a non-UE environment.

However, this weaker characterization fails to shed light on the context dependence of NPIs in non-monotone environments. In the remainder of the section, we explore a characterization of the NPI Licensing Condition that builds on a proposal by von Fintel (1999) and is weaker than the condition in (17) but stronger than the condition in (81).

In characterizing the new NPI licensing condition we first provide a definition of what we will call *Strawson DE environments*, which is a weakening of downward-entailingness above with respect to a position in a constituent, as defined in (5). It is based on the notion of Strawson Entailment, which is entailment on the assumption of definedness of the conclusion (cf. Gajewski 2011):

(82) *Cross-categorial Strawson Entailment* ($\subseteq_S$)

a. For p, q of type t: p $\subseteq_S$ q iff p = 0 or q = 1.

b. For f, g of type $\langle \delta, \tau \rangle$: f $\subseteq_S$ g iff for all x of type δ such that g(x) is defined: f(x) $\subseteq_S$ g(x).

This notion allows us to define Strawson downward-entailingness, as in (83), as well as Strawson downward-entailingness with respect to a position in a constituent, as in (84) (cf. Gajewski 2011, Homer 2012). We say that an element is in a Strawson DE environment if and only if it is in a position with respect to which some constituent in the sentence is Strawson DE.

(83) *Strawson DE function:* A function of type $\langle \delta,\tau \rangle$ is Strawson DE iff for all x, y of type δ such that x $\subseteq_S$ y: f(y) $\subseteq_S$ f(x).

(84) *Strawson DE environment:* A constituent **A** is Strawson DE with respect to α of type δ iff λx. $[\![\mathbf{A}[\alpha/\mathbf{v}_\delta]]\!]^{g[v_\delta \to x]}$ is Strawson DE.

With this in hand, we can now introduce a new NPI licensing condition:

(85) *Modified NPI Licensing Condition* (version 2 of 2)
An NPI is acceptable only if it occurs in a Strawson DE environment.

This condition is weaker than the condition in (17) since not every Strawson DE environment is a DE environment. For instance, if an NPI occurs in the immediate scope of *only John*, which triggers the presupposition that its predicate argument is true of John, it is in a Strawson DE but not a strict DE environment.

(86) Only John read any book.

This is demonstrated in (87) below: if the sentence in the premise is true and the DP in the premise, *any book*, is replaced by a stronger element, say, *several books*, and the meaning of the resulting sentence is defined, then the resulting sentence is true as well. This shows that sentence (86) is Strawson DE with respect to the NPI. If you take out the definedness condition of the conclusion, (87c), the entailment does not go through. This shows that (86) is not DE with respect to the NPI. Accordingly, since *any book* is in a Strawson DE but not a DE environment in (86), it may be acceptable according to the modified NPI Licensing Condition in (85) but not according to the condition in (17).

(87) a. Only John read any book.
b. $[\![\textbf{several books}]\!]^c \subseteq [\![\textbf{any book}]\!]^c$
c. John read several books.
d. ∴ Only John read several books.

NPIs in the immediate scope of non-monotone quantifiers like *exactly n NP* are contained in a constituent that is Strawson DE with respect to their position if two conditions are satisfied (Danny Fox, p.c.): (i) *exactly* is focused, and (ii) focus triggers the existential presupposition that there is an alternative to the clause containing the focused element that is true (e.g., Geurts and van der Sandt 2004). For example, in (88a) *exactly* is focused and triggers the inference in (88b). This inference is predicted if the existential focus closure of (88a) is presupposed and no alternative to *exactly* is DE.

(88) a. EXACTLY two congressmen read some book last year.
b. Presupposition: that two congressmen read some book last year

Sentence (88a) is Strawson DE with respect to the DP *some book*: it holds that if there are exactly two congressmen who read some book and the presupposition of the conclusion in which *some book* is replaced by a stronger expression is true, e.g., there are two congressmen who read two books, then there are exactly two congressmen who read two books.

(89) a. EXACTLY two congressmen read some book.
b. Two congressmen read two books.
c. ∴ EXACTLY two congressmen read two books.

Accordingly, NPIs *may* be acceptable in the scope of non-monotone quantifiers in which *exactly* is focused according to the characterization in (85): they are located in a constituent that is Strawson DE with respect to them. But this is not enough for them to actually *be* acceptable: they are acceptable only if the existential import of focus on *exactly* is satisfied in the context or can be easily accommodated. The existential import of focused *exactly* thus provides a nail on which to hang an account of the context dependence of NPIs under non-monotone quantifiers. However, in the following, we show that it is inadequate.

We focus on the infelicitous sentences in (90). As extensively discussed in Sect. 4, their markedness is predicted by an approach that assumes that the sentences contain a matrix *even* at LF that associates with the NPIs. A matrix *even* would in these examples trigger incorrect scalar presuppositions: given our shared assumptions, it is false that it is less likely that exactly four hundred congressmen read a book last year than, say, that exactly four hundred congressmen read two books; and it is false that it is less likely that exactly eleven of the seventeen players on our soccer team (including substitutes) had contact with the ball rather than, say, that exactly eleven of our players had two contacts with the ball.

(90) a. #EXACTLY four hundred congressmen read any book last year.
b. #EXACTLY eleven players on our soccer team had any contact with the ball.

On the approach described in this section, the infelicity of these sentences could only follow from the existential presuppositions triggered by focus on *exactly* not being satisfied and being difficult to accommodate – otherwise the NPIs would be contained in sentences whose presuppositions are satisfied and that are Strawson DE with respect to them. However, the existential presuppositions triggered by focus on *exactly* in (90) are uncontroversial: as discussed in Sect. 4, it is natural to assume that almost all – say, four hundred – congressmen read a book last year. Similarly, it is practically a given that the eleven starting players on our soccer team had contact with the ball, as well as some substitutes. Thus, it appears not to be possible to ascribe the markedness of the sentences in (90) to a faulty existential presupposition, leaving their infelicity unexplained on the approach described in this section.

5 A general theory of NPI licensing

We have resolved the first two challenges described in the Introduction. In particular, we have shown that a uniform treatment of weak *even* and NPIs in the scope of non-monotone quantifiers allows us to capture the observed context dependence of NPIs in these environments.

(20) *Summary of the challenges*
(I) Distribution of weak *even* under non-monotone quantifiers ✓
(II) Distribution of NPIs under non-monotone quantifiers ✓
(III) Compatibility with a general theory of NPI licensing

In the remainder of the paper, we explore how our proposal relates to two comprehensive theories of NPI licensing. The first theory, already mentioned in the Introduction and restated below, is a trivial generalization of our proposal from the preceding section and takes NPIs to be weak associates of covert *even* tout court (cf. Lee and Horn 1994, Lahiri 1998).

(18) *Single operator theory of NPI licensing*
NPIs are weak elements that are associates of *even*.

Several arguments have been put forward against such a theory – in particular, arguments pertaining to distinct distributions of weak *even* and NPIs in certain environments that are not strictly DE (e.g., Heim 1984, Lahiri 1998, Schwarz 2000). We revisit these arguments and conclude that the single operator theory of NPI licensing is more viable than previously thought.

The second theory, also previewed in the Introduction, takes NPIs to be weak associates of either covert *even* or a covert exhaustification operator *exh* (cf. Krifka 1995, Chierchia 2013). We show that our proposal about NPIs under non-monotone quantifiers can be adequately embedded into this theory once the theory is slightly amended (e.g., it needs to be assumed that all NPIs can be associates of covert *even*) and once more intricate aspects of it are taken into account (intervention effects).

(19) *Multiple operator theory of NPI licensing*
NPIs are weak elements that are associates of either *even* or *exh*.

A proper adjudication between these approaches to NPI licensing is beyond the scope of this paper. We content ourselves with showing that the two approaches are on a more equal footing than usually assumed.

5.1 The single operator theory of NPI licensing

The distribution of NPIs in non-monotone and strict DE environments can be straightforwardly captured by assuming that they are weak elements that are associates of *even*: Lahiri (1998) has forcefully shown this for strict DE environments (see Sect. 4.1) and we have extended his analysis to the scope of non-monotone quantifiers (see Sect. 4.2). Another attractive feature of this assumption is that it directly encodes a generalization of Kadmon and Landman's (1993) Strengthening Condition by requiring the structures with NPIs to be less likely than their alternatives (i.e., logically stronger, in case the structures with NPIs and their alternatives stand in an entailment relation).

The theory gives rise to the expectation that, all else being equal, weak *even* and NPIs should have the same distribution. Although this expectation is borne out when it comes to non-monotone and strict DE environments, as we have seen above, their distributions come apart (i) in certain Strawson DE environments that are not strictly DE and (ii) in questions. In contrast to previous assessments, we show that the discrepancies in the distributions of weak *even* and NPIs are not problematic for the single operator theory of NPI licensing. More to the point, the occurrences of NPIs are correctly predicted by the theory to be (i) context-independent in the respective Strawson DE environments and (ii) not to induce negative bias in questions (pace Heim 1984, Lahiri 1998, Schwarz 2000). On the other hand, the behavior of weak *even* in these environments is unexpected given our assumptions about *even*. We outline some potential explanations for this, though we leave their detailed investigation to another occasion.

5.1.1 NPIs in certain Strawson DE environments

Heim (1984) and Schwarz (2000) discuss the distribution of weak *even* and NPIs in restrictors of universal quantifiers and other Strawson DE environments.[14] They observe that weak *even* exhibits nontrivial context dependence in these environments. Consider the sentences in (91): while (91a) is felicitous (intuitively because reading books and passing exams are related in the context), (91b) is infelicitous (intuitively because reading books and wearing blue jeans are unrelated in the context). (Parallel distributional patterns can be reproduced with some other quantifiers, plural definite descriptions, and with conditionals. The reasoning described in the following applies mutatis mutandis to these cases as well.)

(91) a. Everyone who read even ONE book passed the exam.
b. #Everyone who read even ONE book wore blue jeans.

The behavior of NPIs like *any* and *ever* is different: minimal counterparts of the preceding sentences with NPIs instead of weak *even* are felicitous, i.e., NPIs in these environments do not exhibit context-sensitive behavior.

(92) a. Everyone who read any book passed the exam.
b. Everyone who read any book wore blue jeans.

Diagnosis: unexpected behavior of weak 'even'

A uniform treatment of weak *even* and NPIs does not predict this asymmetry. More specifically, the movement approach to *even* predicts the sentences with weak *even* in (91) to all be felicitous in natural contexts, in parallel to the sentences with NPIs. We show this presently.

Recall that on the movement approach to *even* we may assign the sentences in (91) structures along the lines of (93), where *VP* stands for *pass the exam*/*wear blue jeans* and where *even* takes scope above the universal quantifier (to avoid triggering an unsatisfiable presupposition).

[14] If *every* is taken to trigger an existential presupposition (e.g., Barwise and Cooper 1981), its restrictor is a Strawson DE but not a strict DE environment: a substitution of a weaker expression with a stronger expression in the restrictor may yield an undefined meaning if the resulting restrictor is empty (see Sect. 4.3). On the other hand, if *every* does not trigger an existential presupposition, as recently put forward by Schlenker (2012, Sect. 4.2), its restrictor is a strict DE environment.

(93) **[even$_{\mathbf{C}}$ [everyone who read one$_{\mathbf{F}}$ book VP]]**

The presuppositions that *even* triggers in such structures are consistent. This is because none of the alternatives in the domain of *even* in (93) distinct from the prejacent entails the prejacent (they are either all entailed by the prejacent or they are mutually logically independent, see footnote 14).

(94) C $\subseteq$ {that everyone who read n books VP: n$\geq$1}

Moreover, it holds that in natural contexts in which some people have read several books, the prejacent of *even* in (93) contextually entails what may count as all the relevant alternatives. Namely, if everyone who read one book passed the exam/wore blue jeans and there are people who read several books, then it follows that everyone who read several books passed the exam/wore blue jeans. Consequently, the scalar presuppositions triggered by the sentences in (91) on the construal in (93), represented below, are easily satisfied.

(95) a. For all relevant n>1: that everyone who read one book passed the exam < that everyone who read n books passed the exam.
b. For all relevant n>1: that everyone who read one book wore blue jeans < that everyone who read n books wore blue jeans.

Specifically, the presuppositions are satisfied if, say, it is possible in the context that someone who read exactly one book failed the exam/did not wear blue jeans, respectively, which are possibilities that may well be compatible with our shared assumptions; in case these possibilities obtain, the prejacents contextually entail but are not contextually equivalent to the alternatives and are thus less likely than them (this is assuming that possibility in a context implies a non-zero degree of likelihood in that context). This means that both sentences in (91) are predicted to trigger scalar presuppositions that are satisfied in contexts compatible with our shared assumptions, and both should thus be felicitous. This is a problem for the approach to *even* adopted in this paper since, in actual fact, only one of the two sentences is felicitous.

(96) *Diagnosis (weak 'even'):* The context-dependent distribution of weak *even* in restrictors of universal quantifiers is unexpected on the movement approach to *even.*

In contrast, the distribution of NPIs in restrictors of universal quantifiers, exemplified in (92), is correctly predicted on the assumption that they are weak associates of covert *even.* As we have just seen, in contexts in which it is part of the common ground that there are people who read several books, the prejacents of *even* in the sentences in (92) on the construal in (97) contextually entail what may count as all the relevant alternatives.

(97) **[even$_{\mathbf{C}}$ [everyone who read any book VP]]**

On this construal, if everyone who read any book passed the exam/wore blue jeans and there are people who read several books, then everyone who read several books passed the exam/wore blue jeans. Moreover, if it is possible in the context that someone who read exactly one book failed the exam/did not wear blue jeans, respectively, the scalar presuppositions are satisfied in the context: the prejacents contextually entail but are not contextually equivalent to the alternatives and may thus well be less likely than them. The scalar presuppositions of the sentences in (92), which are equivalent to those in (95), are thus satisfied and so the NPIs contained in them should be licensed. This prediction is borne out.

(98) *Diagnosis (NPIs):* The context-independent distribution of NPIs in restrictors of universal quantifiers is expected on the single operator theory of NPI licensing.

Thus, the distribution of NPIs in restrictors of universal quantifiers is not problematic for the single operator theory of NPI licensing: their apparent context independence is correctly predicted by the theory (pace Heim 1984, Schwarz 2000). However, the context-dependent distribution of weak *even* is unexpected on the movement approach to *even*. In the following, we outline how the puzzling distribution of weak *even* might be explained.

A suggestive parallelism

Consider once more the distribution of *even* in (91), repeated below.

(91) a. Everyone who read even ONE book passed the exam.
b. #Everyone who read even ONE book wore blue jeans.

This ill-understood pattern is paralleled by the pattern in (99), where matrix *even* associates with a weak element that is modified by the exhaustification expression *exactly* (the same pattern obtains if *exactly* is replaced with *only* or *just*). (This parallelism can also be observed with some other quantifiers, with plural definite descriptions, and with conditionals.)

(99) a. Even everyone who read exactly ONE book passed the exam.
b. #Even everyone who read exactly ONE book wore blue jeans.

We have a good handle on the contrast in (99). On the one hand, the scalar presupposition triggered in (99a), given in (100), is satisfied in contexts in which the more books you read, the better you do in exams. On the other hand, the scalar presupposition triggered in (99b), given in (101), is satisfied in contexts in which the more books you read, the more likely you are to wear blue jeans. Accordingly, the difference in felicity between the two sentences is predicted: while the scalar presupposition of the first sentence is satisfied in contexts compatible with our shared assumptions, the scalar presupposition of the second sentence is not.

(100) For all relevant $n>1$: that everyone who read exactly one book passed the exam $<$ that everyone who read exactly n books passed the exam.

(101) For all relevant n>1: that everyone who read exactly one book wore blue jeans < that everyone who read exactly n books wore blue jeans.

We put forward that the facts observed by Heim and Schwarz pertaining to the context dependence of weak *even* in restrictors of universal quantifiers follow from the sentences having the same meanings as the sentences in (99). That is, we submit that the embedded clauses in (91) have a stronger meaning than assumed above: they are covertly exhaustified.

Embedded exhaustification and weak 'even'

Many expressions in natural language give rise to scalar implicatures. An example of a scalar implicature is the inference that usually accompanies the sentence in (102) – that John did not read more than one book.

(102) John read one book ($\rightsquigarrow$ John did not read two books)

There are two types of approaches to deriving scalar implicatures: a pragmatic approach and a grammatical approach. On the grammatical approach, there is a covert operator, *exh*, in grammar that generates scalar implicatures (Krifka 1995, Landman 1998, Fox 2007, Chierchia et al 2011, among others). The meaning of the operator is close to that of *only*: it conveys that the prejacent of *exh* is true and that all the relevant alternatives that are not entailed by the prejacent are false (see Fox 2007 for discussion). The alternatives are determined by the scalar items in the scope of *exh*.

(103) $[\![\mathbf{exh_C}]\!]^c$(p, w) = 1 iff p(w) = 1 & ∀q∈C [p⊈q → q(w)=0]

To illustrate: on this approach to scalar implicatures the sentence in (102) has the structure in (104), where *exh* associates with *one* and yields the meaning that John read one book but not more than one book.

(104) a. **[exh$_{\mathbf{C}}$ [John read one book]]**
b. $[\![$(104a)$]\!]^c$(w) = 1 iff John read one book in w & ∀q∈C ⊆ {that John read n books: n≥1}: [that John read one book ⊈ q → q(w)=0] iff John read exactly one book in w.

If the sentences in (91) are parsed with an embedded *exh* associating with the focused element *one* (we assume that *exh* does not use up the alternatives activated in its scope; see Crnič 2013a for discussion), the resulting structure is thus:

(105) **[even$_{\mathbf{C}}$ [everyone$_{\mathbf{x}}$ [exh$_{\mathbf{C'}}$ [x read one$_{\mathbf{F}}$ book]] VP]]**

The meaning of the sentences and the alternatives they induce will then correspond to those of the sentences in (99): the prejacent of *even* conveys that the VP holds of everyone who read exactly one book, while the alternatives convey that the VP holds of everyone who read exactly n books, where the VP stands for *pass the exam* and *wear blue jeans*, respectively.

(106) a. $[\![$[**everyone$_{\mathbf{x}}$** [**exh$_{\mathbf{C'}}$** [**x read one$_{\mathbf{F}}$ book**]] **VP**]$]\!]^c$(w) = 1 iff everyone who read exactly one book VP-ed in w.
b. ALT([**everyone$_{\mathbf{x}}$** [**exh$_{\mathbf{C'}}$** [**x read one$_{\mathbf{F}}$ book**]] **VP**]) = {that everyone who read exactly n books VP-ed: n≥1}

Accordingly, the scalar presupposition triggered by the sentence with the VP *pass the exam*, which is identical to that of the felicitous sentence in (99), will be satisfied in contexts compatible with our shared assumptions, i.e., contexts in which the more books you read, the better you do in exams:

(107) a. Everyone who read even ONE book passed the exam.
b. [**even$_{\mathbf{C}}$** [**everyone$_{\mathbf{x}}$** [**exh$_{\mathbf{C'}}$** [**x read one$_{\mathbf{F}}$ book**]] **passed the exam**]]
c. $[\![$(107b)$]\!]^c$ is defined only if for all relevant n>1: that everyone who read exactly one book passed the exam < that everyone who read exactly n books passed the exam.

In contrast, the scalar presupposition triggered by the sentence with the VP *wear blue jeans*, which is identical to that of the infelicitous sentence in (99), will not be satisfied in contexts compatible with our shared assumptions.

(108) a. Everyone who read even ONE book wore blue jeans.
b. [**even$_{\mathbf{C}}$** [**everyone$_{\mathbf{x}}$** [**exh$_{\mathbf{C'}}$** [**x read one$_{\mathbf{F}}$ book**]] **wore blue jeans**]]
c. $[\![$(108b)$]\!]^c$ is defined only if for all relevant n>1: that everyone who read exactly one book wore blue jeans < that everyone who read exactly n books wore blue jeans.

By assuming that the sentences in (91) are parsed with embedded *exh* we are thus able to derive the otherwise puzzling contrast in their felicity.

(109) *Descriptive generalization:* The distribution of weak *even* in restrictors of universal quantifiers is explained on the assumption that the associate of *even* is exhaustified in this environment.

Two questions are raised by this proposal: Why should restrictors of *every* containing weak *even* be exhaustified? And what is the scope of interaction between *even* and *exh*? We will return to these questions briefly in the conclusion.

5.1.2 NPIs in questions

It is well known that questions containing weak *even* effect negative bias – the inference that the speaker expects a negative answer to the question (Borkin 1971, Heim 1984, van Rooy 2003, Abels 2003, Guerzoni 2003, 2004, among others). For example, the speaker who uses the question in (110) is taken to expect that John doesn't have any friends.

(110) Does John have even ONE friend? (negative bias)

This is not the case if weak *even* is replaced with an NPI (though see Han and Siegel 1997 for caveats): the sentence in (111) effects no negative expectation concerning the speaker's information state. This is another distributional difference between weak *even* and NPIs that on the face of it appears to argue against the single operator theory of NPI licensing.

(111) Does John have any friends? (no bias)

In the following we show that on certain approaches to questions (e.g., Nicolae 2013, Guerzoni and Sharvit 2013), the single operator theory of NPI licensing correctly predicts NPIs to be licensed in questions and not to induce negative bias. However, once again the fact that weak *even* induces negative bias turns out to be unexpected.

Downward-entailing environments in questions

There are approaches to questions that take them to contain a (Strawson) DE environment (e.g., Nicolae 2013, Guerzoni and Sharvit 2013). For example, Guerzoni and Sharvit (2013) propose that a yes/no-question like (112a) can have the underlying structure in (112b), which contains disjunction and negation (these are optionally pronounced, hence the parentheses); the first disjunct is elided under semantic identity with material in the second disjunct. (We significantly simplify Guerzoni and Sharvit's proposal in the following; the reader is referred to their paper for details. The choice to use Guerzoni and Sharvit's proposal is not crucial, and analogous reasoning extends to, say, the proposal of Nicolae.)

(112) a. Does John have one friend?
b. [**whether$_7$** [~~**J. has one friend**~~ [(**or$_7$** [**not**) **J. has one friend**]]]]

The elements in the structure in (112) compose to yield the meaning that corresponds to the Hamblin set of the question consisting of the affirmative and the negative answer to the question, i.e., the two disjuncts in (112b) (more precisely, *whether* is an existential quantifier over true propositions that are identical either to the first disjunct or to the second disjunct).

(113) $[\![(112b)]\!]^c(w) = \{[\![\textbf{John has one friend}]\!]^c, \neg[\![\textbf{John has one friend}]\!]^c\}$

Adopting these assumptions about questions, it can be easily shown why NPIs are predicted to be acceptable in questions and not to trigger negative bias on the single operator theory of NPIs. For example, the sentence in (111) may have a representation in which *even* scopes above negation, while the NPI remains in situ, as shown in (114). The ellipsis of the first disjunct is licensed since semantic identity obtains between the disjunct and a constituent in the scope of negation in the second disjunct (recall that NPIs denote existential quantifiers).

(114) a. Does John have any friends?
b. **[whether$_7$ [~~John has some friends~~ (or$_7$ [even$_C$ [not) John has any friends]]]]**

The meaning of this structure consist of the affirmative and the negative answer to the question, only the latter of which is presuppositional.

(115) $[\![(114b)]\!]^c = \{[\![\textbf{John has some friends}]\!]^c, [\![\textbf{even}_\textbf{C}]\!]^c(\neg[\![\textbf{John has any friends}]\!]^c)\}$

The scalar presupposition that *even* triggers in the negative answer in (115) is almost trivial since the prejacent of *even* is stronger than all the relevant alternatives and may thus well be less likely than them.

(116) $[\![\textbf{even}_\textbf{C}]\!]^c(\neg[\![\textbf{John has any friends}]\!]^c)$ is defined only if for all n>1: that John doesn't have one friend < that John doesn't have n friends.

For ease of presentation, we assume that presuppositions of Hamblin alternatives project universally (see Fox 2010, Heim 2012 for discussion). Accordingly, the question in (114) inherits a scalar presupposition that is almost trivial and hence does not give rise to negative bias.

(117) *Diagnosis (NPIs):* If questions are analyzed as containing a DE environment, the single operator theory makes the correct prediction that NPIs may be licensed in them and not induce negative bias.

Negative bias in questions

The movement approach to *even* gives rise to the same prediction about the behavior of weak *even* in questions as the single operator theory does about the behavior of NPIs.[15]

[15] Guerzoni (2004) develops an ingenious approach to deriving negative bias in questions. However, on her approach, NPIs are incorrectly predicted to induce negative bias on all operator-based theories of NPIs, as we discuss in the following. Guerzoni argues that the negative bias in (110) emerges from an interaction of *even* and a specific element in the structure of questions, which we will call X for brevity. If *even* stays in situ, the interpretation of the structure is a Hamblin set containing two alternatives – the affirmative answer and the negative answer – both of which contain *even* that associates with a weak element in its immediate scope, making them inconsistent and the question illicit. However, if *even* moves above X at LF, the interpretation of the resulting structure is a Hamblin set containing the affirmative answer in which *even* associates with a weak element in its immediate scope and the negative answer in which negation intervenes between *even* and its weak associate, (i). While the former answer is inconsistent, the latter answer is consistent. This explains the negative bias induced by weak *even* in questions: only the negative answer is assertable. See Guerzoni (2004) for details.

(i) $[\![(110)]\!]^c(\mathrm{w}) = \{\#[\![\textbf{even}_\textbf{C}]\!]^c([\![\textbf{John has one friend}]\!]^c), \checkmark[\![\textbf{even}]\!]^c(\neg[\![\textbf{John has one friend}]\!]^c)\}$

Given this derivation, NPIs are predicted to induce the same bias as weak *even* on all operator-based theories of NPI licensing. These theories are geared towards deriving infelicity

(118) *Diagnosis (weak 'even'):* If questions are analyzed as containing a DE environment, the movement approach to *even* makes the false prediction that weak *even* might not induce negative bias in them.

Namely, if *even* moves above the DE operator in a question, it should trigger a scalar presupposition that is almost trivial. This is illustrated in (119)-(120), where the presupposition triggered by *even* in the negative answer corresponds to the presupposition in (116) and hence cannot be taken to be the source of the negative bias that weak *even* induces.

(119) a. Does John have even ONE friend?
b. **[whether ~~CP~~ (or [even$_{\mathbf{C}}$ [not) [John has one$_{\mathbf{F}}$ friend]]]]**

(120) $[\![$(119b)$]\!]^c = \{[\![\textbf{John has one friend}]\!]^c,$
$[\![\textbf{even}_{\mathbf{C}}]\!]^c(\neg[\![\textbf{John has one}_{\mathbf{F}}\textbf{ friend}]\!]^c)\}$

A further mechanism is thus needed to derive negative bias induced by weak *even*. One candidate for such a mechanism is embedded exhaustification, as entertained in our above discussion. If weak *even* were obligatorily accompanied by embedded exhaustification in questions, the appropriate structure would be as follows:

(121) a. Does John have even ONE friend?
b. **[whether ~~CP~~ (or [even$_{\mathbf{C}}$ [not) [exh$_{\mathbf{C'}}$ [J. has one$_{\mathbf{F}}$ friend]]]]]**

(122) $[\![$(121b)$]\!]^c = \{[\![\textbf{John has one friend}]\!]^c,$
$[\![\textbf{even}_{\mathbf{C}}]\!]^c(\neg[\![\textbf{exh}_{\mathbf{C'}}\textbf{ [John has one}_{\mathbf{F}}\textbf{ friend]}]\!]^c)\}$

Here the scalar presupposition triggered by *even* would be contingent: it is less likely that John does not have exactly one friend than, say, that he does not have exactly two friends.

(123) $[\![\textbf{even}_{\mathbf{C}}]\!]^c(\neg[\![\textbf{exh}_{\mathbf{C'}}\textbf{ [John has one}_{\mathbf{F}}\textbf{ friend]}]\!]^c)$ is defined only if for all n>1: that John does not have exactly one friend < that John does not have exactly n friends.

---

of NPIs in UE environments, which usually follows from an inconsistent inference triggered by the NPI licensing operator, OP. Accordingly, OP cannot be interpreted in situ in questions but must be interpreted above X. The resulting structures have meanings along the lines of (ii), where only the negative answer has a consistent meaning. The questions should thus induce negative bias.

(ii) $[\![$(111)$]\!]^c = \{\#[\![\textbf{OP}_{\mathbf{C}}]\!]^c([\![\textbf{John has any friends}]\!]^c),$
$\checkmark[\![\textbf{OP}_{\mathbf{C}}]\!]^c(\neg[\![\textbf{John has any friends}]\!]^c)\}$

This presupposition entails that John is more likely to have a lower number of friends than a higher number of friends. Although this inference might indeed accompany occurrences of weak *even* in questions, it does not quite correspond to the negative bias that weak *even* induces – that the speaker does not expect the affirmative answer to the question. Accordingly, relying on embedded exhaustification alone does not suffice to derive the negative bias induced by weak *even*. Further research is needed to resolve the puzzle of negative bias.

To summarize: we have shown that the single operator theory of NPI licensing, on which NPIs are effectively treated as weak *even* expressions, correctly predicts the distribution of NPIs across a variety of environments, including their context independence in restrictors of universal quantifiers as well as their acceptability and non-biasing nature in questions. Accordingly, the single operator theory of NPI licensing can be seen as a viable candidate for a general theory of NPIs. On the other hand, the behavior of weak *even* in certain environments turns out to be unexpected on the movement approach to *even*: it exhibits context dependence in restrictors of universal quantifiers and it induces negative bias in questions. We have indicated some possible explanations of this behavior, though much has been left for future research. In the remainder of this section, we now turn to a different theory of NPI licensing and show how our treatment of NPIs in non-monotone environments can be absorbed into it.

5.2 The multiple operator theory of NPI licensing

Building on the work of Krifka (1995), Chierchia (2013) proposes that NPIs are weak elements that can be associates of different alternative-sensitive operators; in particular, some NPIs are associates of covert *even*, while some other NPIs are associates of a covert exhaustification operator *exh*. (For presentational reasons we here simplify various aspects of Chierchia's system; the reader is referred to his manuscript for details.)

Which operator accompanies which NPI depends on the lexical specification of the NPI. Chierchia proposes that the lexical specification goes hand in hand with what alternatives the NPI induces. These alternatives come in two varieties: domain alternatives and scalar alternatives. NPIs like *any* and *ever* induce domain alternatives. Domain alternatives of an NPI are existential quantifiers that differ from the NPI only in that their domain is a subset of the domain of the NPI.

(124) *Domain alternatives of 'any'*
$\mathrm{ALT}(\mathbf{any_D}) = \{\lambda P.\lambda Q.\ \exists x{\in}D'\ [P(x){=}Q(x){=}1]{:}\ D'{\subseteq}D\}$

Chierchia takes the NPIs that induce domain alternatives to be associates of *exh*. This can be implemented in different ways, say, by the NPIs bearing an uninterpretable feature that must be checked by *exh*.

(125) *Association with 'exh'*
NPIs that induce domain alternatives are associates of *exh*.

NPIs that induce (only) scalar alternatives are taken to be associates of covert *even*, which can be implemented by the NPIs bearing an uninterpretable feature that must be checked by *even*. A prominent class of such NPIs are minimizers like *lift a finger* and *a single book* (cf. Heim 1984, Guerzoni 2004).

(126) *Scalar alternatives of 'a single book'*
ALT(**a single book**) = {one book, two books, three books, ...}

(127) *Association with covert 'even'*
NPIs that induce (only) scalar alternatives are associates of *even*.

NPIs like *any* and *ever* are acceptable only if the inference triggered by *exh* that associates with them is consistent. If an NPI is unembedded, this inference will in general be inconsistent. This is illustrated in (128): the sentence conveys that someone read a book in the set D and that they did not read a book in any proper subset of D – which is impossible.[16]

(128) a. #John read any book.
b. **[exh$_{\mathbf{C}}$ [John read any$_{\mathbf{D}}$ book]]**
c. $[\![$(128b)$]\!]^c$(w) = 1 iff John read some book in D in w & ∀q ∈ {that John read some book in D': D'⊆D}: [that John read some book in D ⊈ q → q(w) = 0] iff John read some book in D in w & ∀D'⊂D: John did not read a book in D' in w. (unsatisfiable)

If *any* is embedded under a DE operator and *exh* takes matrix scope, the inference triggered by *exh* will be vacuous – and thus consistent – since the prejacent of *exh* entails all of the alternatives. For example, if John did not read a book in D, then he did not read a book in any subset D' of D.

(129) a. John didn't read any book.
b. **[exh$_{\mathbf{C}}$ [not [John read any$_{\mathbf{D}}$ book]]]**
c. $[\![$(129b)$]\!]^c$(w) = 1 iff John did not read a book in D in w.

[16] This reasoning would not go through if one were able to prune domain alternatives over which *exh* quantifies, i.e., if the domain of *exh* could be restricted to a proper subset of the alternatives induced by an NPI. For example, this would be the case if the prejacent in (128) would convey that John read a book written by Tolstoy, Dostoevsky, or Flaubert, and the domain alternatives would be restricted to 'that John read a book written by Tolstoy' and 'that John read a book written by Dostoevsky'; exhaustifying the sentence with respect to these two alternatives would yield a consistent meaning that John read a book by Flaubert. To avoid this issue, a ban on pruning of domain alternatives could be stated as a primitive of the theory. However, such a ban likely follows from more general constraints on pruning of alternatives. For example, Fox and Katzir (2011) characterize a restriction on pruning that appears to encompass the ban on pruning of domain alternatives.

5.2.1 An apparent issue with non-monotone environments

All else being equal, an approach that takes *any* and *ever* to induce domain alternatives and to be associates of *exh* fails to correctly predict their context dependence in non-monotone environments. We show this on the basis of a concrete example. If *any* in the infelicitous sentence in (130a) is accompanied by *exh*, the sentence may have the structure in (130b) and convey that exactly four hundred congressmen voted for a proposition in D this week (= the prejacent of *exh*), but that for any subset D' of D, it is false that exactly four hundred congressmen voted for a proposition in D' (= the exhaustive inference triggered by *exh*). This inference is consistent.

(130) a. #Exactly four hundred congressmen voted for any proposition this week.
b. **[exh$_{\mathbf{C}}$ [exactly 400 congressmen voted for any$_{\mathbf{D}}$ prop]]**
c. $[\![(130\text{b})]\!]^c(\text{w}) = 1$ iff exactly 400 congressmen voted for a proposition in D in w & ∀D'⊂D: ¬(exactly 400 congressmen voted for a proposition in D' in w).

Imagine the following scenario, which is not implausible: two propositions have been brought before Congress so far this week; one proposition had a Republican sponsor, while the other had a Democratic sponsor. Due to the extreme partisanship in Congress, only the Republicans voted for the proposition sponsored by the Republican and only the Democrats voted for the proposition sponsored by the Democrat (and thirty-five congressmen were absent). The sentence in (130) describes such a scenario: it holds that exactly four hundred congressmen voted for one of the two propositions but less than four hundred congressmen voted for the Republican-sponsored proposition and less than four hundred congressmen voted for the Democrat-sponsored proposition. Since the sentence is perceived as infelicitous even though it can describe such plausible scenarios, the multiple operator theory of NPIs that treats *any* as an associate of *exh* overgenerates. However, this overgeneration is curtailed once intervention is taken into account.

5.2.2 Minimality and intervention

A similar issue with overgeneration arises elsewhere. Since the theory predicts that *any* and *ever* are unacceptable only if the exhaustive inference triggered by *exh* that associates with them is inconsistent, the theory appears to falsely predict that NPIs may be acceptable in certain UE environments. One such environment is the scope of a universal quantifier (cf. Chierchia 2013):

(131) #Every student read any book.

The sentence in (131) may be assigned a structure in which *exh* associates with the NPI across the universal quantifier. The meaning of such a configuration is that every student read a book in D but not every student read a book in some proper subset of D:

(132) a. **[exh$_C$ [every student read any$_D$ book]]**
b. ⟦(132a)⟧^{c}(w) = 1 iff every student read a book in D in w & ∀D'⊂D: ¬(every student read a book in D' in w).

This meaning is consistent. For instance, it may describe a situation in which three students read three different books, each reading just one book: it is true that every student read one of the three books but it is false that every student read a book in a set that is lacking one of the three books. Accordingly, further assumptions are needed to derive the unacceptability of the NPI in (131).

Chierchia proposes that the sentence in (131) is ruled out once one takes into account that the relation between *exh* and its associate is subject to Minimality (a principle that requires operators to affect their closest targets; Rizzi 1990). Specifically, if *exh* exhaustifies the domain alternatives of *any*, it must also exhaustify the domain alternatives of all intervening quantifiers, i.e., of all quantifiers that are asymmetrically c-commanded by *exh* and which in turn c-command *any*. Accordingly, *exh* associates with multiple elements in sentence (131):

(133) **[exh$_C$ [everyone$_D$ read any$_B$ book]]**

The domain of *exh* in (133) thus consists of alternatives that differ from the prejacent in that *every* and *any* quantify over subsets of the domains of *every* and *any* in the prejacent. The inference generated by *exh* in (133), represented in (134) below, is inconsistent given these alternatives: it cannot be the case that every student in D read a book in B but that there are no subsets D' and B' of D and B (other than D and B themselves) such that every student in D' read a book in B'.

(134) ⟦(133)⟧^{c}(w) = 1 iff every student in D read a book in B in w & ∀D'⊆D, B'⊆B: if (D',B') ≠ (D,B), then ¬(every student in D' read a book in B' in w). (unsatisfiable)

This means that once Minimality is taken into account, a sentence like (131) necessarily gives rise to an inconsistent meaning, explaining the unacceptability of the NPI in the immediate scope of a universal quantifier.

5.2.3 Back to NPIs in non-monotone environments

If previously the multiple operator theory faced the problem of overgeneration, it now faces the problem of undergeneration: due to Minimality, the prediction is that NPIs should be as unacceptable in the scope of non-monotone quantifiers as they are in UE environments, no matter what the shared contextual assumptions are. This is exemplified in (135)-(136): due to the association pattern in (135), the exhaustive inference triggered by *exh*, computed in (136), is inconsistent.[17]

(135) a. Exactly two congressmen read any book last year.

b. [**exh**$_{\mathbf{C}}$ [**exactly two**$_{\mathbf{D}}$ **congressmen read any**$_{\mathbf{B}}$ **book**]]

(136) $[\![(135b)]\!]^c$(w) = 1 iff exactly two congressmen in D read a book in B in w & ∀D'⊆D, B'⊆B: if (D',B',)≠(D,B), then ¬(exactly two congressmen in D' read a book in B' in w). (unsatisfiable)

This issue of undergeneration can be avoided by making a minor amend-mendment to the multiple operator theory: while NPIs like *any* and *ever* may induce domain alternatives, as assumed by Chierchia and Krifka, they may also induce just scalar alternatives, as we assumed in the preceding section (cf. Lee and Horn 1994, Lahiri 1998). This means that NPIs like *any* and *ever* are effectively ambiguous. In line with Chierchia's proposal we can take them to be associates of *exh* when they induce domain alternatives and associates of covert *even* when they induce scalar alternatives.

(137) *Domain alternatives of 'any' (association with 'exh')*
ALT(**any**$_{\mathbf{D}}$) = {λP.λQ. ∃x∈D' [P(x)=Q(x)=1]: D'⊆D}

(138) *Scalar alternatives of 'any' (association with 'even')*
ALT(**any**$_{\mathbf{D}}$) = {one, two, three, ...}

We have just seen that if NPIs in the scope of non-monotone quantifiers are construed as inducing domain alternatives and are accompanied by *exh*, the sentences have inconsistent meanings. So, NPIs cannot be construed as inducing domain alternatives in the scope of non-monotone quantifiers. The remaining option is to construe them as inducing scalar alternatives.

If NPIs are construed as inducing scalar alternatives and are accompanied by *even*, the scalar presupposition that *even* triggers may be satisfied. We have discussed this extensively in Sect. 4. In the following we show that Minimality does not significantly affect the conclusions reached there. For example, in a sentence like (139), in which *even* associates with an NPI that induces scalar alternatives, *even* must also associate with all intervening scalar items, due to Minimality.

(139) a. Exactly two congressmen read any book last year.

b. [**even**$_{\mathbf{C}}$ [**exactly two congressmen read any book**]]

[17] The exhaustive inference in (136) is inconsistent only on the perhaps uncontroversial assumption that the restrictor of the non-monotone quantifier *exactly n NP* must contain more than n individuals, which may well be an antipresupposition of the quantifier.

The domain of *even* thus consists of alternatives that have the form in (140).

(140) C $\subseteq$ {that exactly m congressmen read n books last year: m,n$\geq$1}

The scalar presupposition triggered by *even* in (139) is satisfied in a context compatible with our assumptions if the relevant alternatives are, say, that exactly four hundred congressmen read a book, which conforms to our shared assumption that about four hundred congressmen read a book and is thus quite likely, and that exactly two congressmen read five/ten books, etc. (see the preceding section for detailed discussion).

(141) $[\![$(139b)$]\!]^c$ is defined only if for all relevant m,n$\geq$1: if (m,n)$\neq$(2,1), then that exactly two congressmen read a book last year $<$ that exactly m congressmen read n books last year.

Finally, the fact that *any* and *ever* may induce scalar alternatives that are used up by *even* does not cause problems with respect to the distribution of NPIs in UE and DE environments. For example, if an NPI occurs in an UE environment, at least some of the relevant alternatives it induces will be logically stronger and, accordingly, at most as likely as the prejacent, as illustrated in (142). The scalar presupposition triggered by *even* will thus be unsatisfiable and the NPIs unacceptable.

(142)
a. #Every student read any book
b. **[even$_{\mathbf{C}}$ [every student read any book]]**
c. $[\![$(142b)$]\!]^c$(w) is defined only if for all relevant n>1: that every student read a book $<$ that every student read n books, that some student read n books. (unsatisfiable)

To summarize: we have shown that although the distribution of NPIs in the scope of non-monotone quantifiers at first appears to be problematic for the multiple operator theory of NPI licensing, a more comprehensive and slightly amended version of the theory – a version that takes into account Minimality and that assumes that NPIs like *any* and *ever* may induce scalar alternatives that are used up by *even* – allows us to adequately explain the distribution of NPIs in non-monotone environments, as well as in environments previously discussed by Krifka (1995) and Chierchia (2013).

We have thus resolved the third and final challenge of the paper: the proposal that we put forward about NPI licensing under non-monotone quantifiers is compatible with a general theory of NPI licensing. This is trivially the case for the single operator theory, which we have shown to be a more viable theory of NPI licensing than commonly assumed, but it also holds for the multiple operator theory.

(20) *Summary of the challenges*
(I) Distribution of weak *even* under non-monotone quantifiers ✓
(II) Distribution of NPIs under non-monotone quantifiers ✓
(III) Compatibility with a general theory of NPI licensing ✓

6 Conclusion

The scalar particle *even* that associates with a weak element in its immediate surface scope, weak *even*, is acceptable only if it is embedded in a non-UE environment, e.g., in the scope of a non-monotone quantifier. This restriction on its distribution is explained by assuming (i) that *even* triggers a scalar presupposition that its prejacent is less likely than all the relevant alternatives and (ii) that it can move at LF (Karttunen and Peters 1979, Wilkinson 1996, Lahiri 1998, among others). If *even* moves above a non-monotone quantifier and strands its weak associate in the scope of that quantifier, its scalar presupposition *may* be licit – but *is* actually licit only if the requisite likelihood relation obtains in the context or can be easily accommodated.

NPIs in non-monotone environments exhibit behavior that mirrors that of weak *even*. Accordingly, we transposed our account of the distribution of weak *even* to NPIs: NPIs in non-monotone environments denote weak elements that are associates of covert *even* (cf. Lee and Horn 1994, Lahiri 1998). We have thus provided a syntactic-semantic explanation of the distribution of NPIs in these environments that eschews informal notions like pragmatic licensing (e.g., Linebarger 1987). Finally, we have shown that our account is compatible with more general theories of NPI licensing.

Avenues for future research have been raised throughout this paper. We conclude by briefly elaborating on three of them. The first avenue concerns the question of how the single operator theory of NPI licensing fits into the more general system of polarity. The second and third avenue concern the interaction of *even* and covert exhaustification.

6.1 Polarity system

There are many issues in the literature on polarity licensing that the single operator theory of NPI licensing should be able to deal with (see Chierchia 2013 for a detailed overview). We outline two of them here, as well as initial strategies of how to go about them. First: Linebarger (1987) observed that NPIs have to stand in a particular relation to a Strawson DE operator to be acceptable in its scope. Chierchia (2004) qualifies Linebarger's observation with respect to scalar items by noting that NPIs are acceptable only if they are not in the immediate scope of a UE strong scalar item like *everyone*.

(143) a. Mary doubts that John ate any donuts.
b. Mary doubts that someone ate any donuts.
c. #Mary doubts that everyone ate any donuts.

To deal with such data in his system, Chierchia (2013) proposes that in sentences like (143), the operator that associates with an NPI obligatorily associates with – and activates the alternatives of – all scalar items that intervene between the operator and the NPI, an assumption that he takes to follow from Minimality. If a similar assumption is made on the single operator theory of

NPI licensing, a contrast between (143b) and (143c) is predicted as well. On this assumption, the alternatives over which *even* quantifies in (143b,c) are the following:

(144) {that Mary doubts that someone ate n donuts, that Mary doubts that everyone ate n donuts: n≥1}

It holds that the prejacent of covert *even* in (143b) – that Mary doubts that someone ate some donuts – entails all the relevant alternatives; this is not the case in (143c), where the prejacent is that Mary doubts that everyone ate some donuts, which is entailed by the alternative that Mary doubts that someone ate some donuts. Accordingly, the scalar presupposition triggered by *even* – that the prejacent is less likely than all the relevant alternatives – may be satisfied in the former case, (143b), but not the latter, (143c).

Second: Another topic that needs to be investigated is the distribution of so-called free choice *any* (e.g., Dayal 1998, Chierchia 2013). An example with a free choice occurrence of *any* is provided in (145a), which conveys that every book is such that John can read it. If the domain of books consists of *Syntactic Structures* and *Inquiry*, as we assume in the following, then the sentence in (145a) has the meaning that John may read either of them, (145b).

(145) a. John can read any book.
b. ⇒ John can read Syntactic Structures & John can read Inquiry.

Since on the single operator theory *any* is a weak associate of *even*, a uniform account of the distribution of *any* would require there to be a covert *even* that associates with the *any* in (145a) as well. We could assume that the free choice inference is generated in grammar (e.g., Fox 2007, Aloni 2007 for free choice disjunction) and that *even* may take scope above the mechanism that generates it, represented with *FC* in the following structure:

(146) **[even$_{\mathbf{C}}$ [FC [John can read any$_{\mathbf{D}}$ book]]]**

On these assumptions, a uniform account would give us the right result if the prejacent of *even* has the meaning described in (147a) and the alternatives in (147b,c).

(147) a. John can read Syntactic Structures & John can read Inquiry.
b. John can read Syntactic Structures.
c. John can read Inquiry.

The scalar presupposition triggered by *even* in (146) may in this case be satisfied since the prejacent of *even*, (147a), would entail all the alternatives and may well be less likely than them. However, to obtain the alternatives in (147) one would need to depart from the simplifying assumption that the alternatives to NPIs are solely number indefinites and switch to the more sophisticated assumption that NPIs may also have existential quantifiers whose domains are subsets of the domain of the NPIs as alternatives, as discussed in

Sect. 5.2 (Krifka 1995, Chierchia 2013). A more serious investigation of these issues as well as other issues pertaining to polarity phenomena constitutes the first avenue for future research.

6.2 Obligatory embedded exhaustification

Weak *even* exhibits nontrivial context dependence in certain environments, e.g., in restrictors of universal quantifiers. We have shown in Sect. 5.1 that this context dependence is unexpected, but that it can be explained on the assumption that the associate of *even* is exhaustified in these environments.

(91) a. Everyone who read even ONE book passed the exam.
b. #Everyone who read even ONE book wore blue jeans.

Interestingly, the environments in which such context dependence can be observed and for which embedded exhaustification has been postulated parallel those that have been independently argued to allow embedded exhaustification in the scope of *even*. Specifically, Crnič (2013a) has argued that the felicity of the sentence in (148) can be explained only on the assumption that the restrictor of the universal quantifier is exhaustified; otherwise the matrix *even* would trigger a scalar presupposition that clashes with the principle of entailment (the strong associate of *even*, which is *all*, is in a Strawson DE environment; see Crnič 2013a for details).

(148) Even everyone who read ALL of the books failed the exam.

Thus, the distributional pattern exemplified in (91) appears to be an instance of a more general phenomenon. In fact, the data discussed in this paper suggest that the interaction between *even* and covert exhaustification might be more systematic than previously assumed, as postulated in (149). The second avenue for future research is to determine the forces governing this interaction.

(149) *Embedded exhaustification generalization:* If covert exhaustification can apply in the scope of (overt) *even*, it must apply.

6.3 Precluded embedded exhaustification

The existence of covert exhaustification in grammar raises a challenge for the approaches to weak *even* and NPIs discussed in this paper. For example, the movement approach to *even* predicts that if weak *even* is generated in the scope of an *exh* that associates with a scalar item that c-commands weak *even*, weak *even* may be acceptable. Thus, if the sentence in (150a) could have the structure in (150b), where *exh* associates with *two*, the sentence should have the same acceptability status as the sentence in (151), contrary to fact.

(150) a. #Two congressmen read the constitution even ONCE.

b. **[even$_{\mathbf{C}}$ [exh$_{\mathbf{D}}$ [two congressmen read the constitution once$_{\mathbf{F}}$]]**

(151) Exactly two congressmen read the constitution even ONCE.

Rothschild (2006) points out a parallel puzzle for NPIs: if *exh* were present in grammar, there should in principle be no difference in the acceptability of the sentences in (152), since they could have comparable structures and semantics.

(152) a. #Two congressmen read any book last year.
b. Exactly two congressmen read any book last year.

We tentatively submit that the inability of covert exhaustification to rescue contradictory occurrences of weak *even* and NPIs reflects a preference to achieve interpretability and assertability in grammar by overt means – if these are available and not weakening relative to their covert counterparts. Clearly, this is by no means a full account of the interaction of *exh* and weak *even*/NPIs, but only a suggestion of a direction in which such an account could be pursued. This pursuit constitutes the third avenue for future research.

Acknowledgements I am grateful to Kai von Fintel, Irene Heim, Roni Katzir, and Daniel Rothschild for their helpful comments. Special thanks to Danny Fox and Gennaro Chierchia. Thanks also to the participants of the Language, Logic and Cognition Seminar at the Hebrew University, to the members of the semantics group at Institut Jean Nicod, to Christine Bartels, and to two anonymous reviewers for *Natural Language Semantics*.

References

Abels K (2003) Who gives a damn about minimizers in questions? In: Young RB, Zhou Y (eds) Proceedings of SALT 13, CLC Publications, Ithaca, NY, pp 1–18

Aloni M (2007) Free choice, modals, and imperatives. Natural Language Semantics 15(1):65–94

Barwise J, Cooper R (1981) Generalized quantifiers and natural language. Linguistics and Philosophy 4(2):159–219

Bennett J (1982) Even if. Linguistics and Philosophy 5(3):403–418

Borkin A (1971) Polarity items in questions. In: Proceedings of CLS, Chicago Linguistic Society, Chicago, vol 7, pp 53–62

Chierchia G (2004) Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In: Belletti A (ed) Structures and beyond, Oxford University Press, pp 39–103

Chierchia G (2013) Logic in Grammar. Oxford University Press, Oxford

Chierchia G, Fox D, Spector B (2011) The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In: Portner P, Maienborn C, von Heusinger K (eds) Handbook of Semantics, vol 2, Mouton de Gruyter, pp 2297—2332

Chomsky N (1976) Conditions on rules of grammar. Linguistic Analysis 2:303–351

Crnič L (2013a) Focus particles and embedded exhaustification. Journal of Semantics 30(4):533–558

Crnič L (2013b) How to get *even* with desires and imperatives. In: Csipak E, Eckardt R, Liu M, Sailer M (eds) Beyond *ever* and *any*, Mouton de Gruyter, Berlin, pp 127–154

Dayal V (1998) *Any* as inherently modal. Linguistics and philosophy 21(5):433–476

Drubig HB (1994) Island constraints and the syntactic nature of focus and association with focus. In: Arbeitspapiere des Sonderforschungsbereichs 340: Sprachtheoretische Grundlagen der Computerlinguistik, University of Tübingen

von Fintel K (1999) NPI licensing, Strawson entailment, and context dependency. Journal of Semantics 16(2):97–148

Fox D (2007) Free choice and the theory of scalar implicatures. In: Sauerland U, Stateva P (eds) Presupposition and Implicature in Compositional Semantics, Palgrave Macmillan, Basingstoke, pp 71–120

Fox D (2010) Lectures on questions. Handout, MIT

Fox D, Katzir R (2011) On the characterization of alternatives. Natural Language Semantics 19(1):87–107

Gajewski J (2005) Neg-raising: Presupposition and polarity. PhD thesis, MIT

Gajewski J (2008) NPI *any* and connected exceptive phrases. Natural Language Semantics 16(1):69–110

Gajewski J (2011) Licensing strong NPIs. Natural Language Semantics 19(2):109–148

Geurts B, van der Sandt R (2004) Interpreting focus. Theoretical Linguistics 30:1–44

Guerzoni E (2003) Why *even* ask? On the pragmatics of questions and the semantics of answers. PhD thesis, MIT

Guerzoni E (2004) *Even*-NPIs in yes/no questions. Natural Language Semantics 12(4):319–343

Guerzoni E, Sharvit Y (2013) Whether or not anything but not whether anything or not. Manuscript, USC and UCLA

Halpern JY (2003) Reasoning about uncertainty. MIT Press, Cambridge, MA

Han CH, Siegel L (1997) Syntactic and semantic conditions on NPI licensing in questions. In: Agbayani B, Tang SW (eds) Proceedings of the WCCFL 15, CSLI Publications, Stanford, CA, pp 177–191

Heim I (1984) A note on negative polarity and downward entailingness. In: Proceedings of NELS, GLSA Publications., Amherst, MA, vol 14, pp 98–107

Heim I (1992) Presupposition projection and the semantics of attitude verbs. Journal of Semantics 9(3):183

Heim I (2012) Functional readings without type-shifted nouns. Manuscript, MIT

Hintikka J (1962) Knowledge and belief: an introduction to the logic of the two notions. Cornell University Press, Ithaca, NY

Homer V (2012) Domains of polarity items, accepted for publication by *Journal of Semantics*

Kadmon N, Landman F (1993) *Any*. Linguistics and Philosophy 16(4):353–422

Karttunen L, Peters S (1979) Conventional implicature. In: Oh CK, Dinneen DA (eds) Syntax and semantics, vol 11, Academic Press, pp 1–56

Kay P (1990) *Even*. Linguistics and Philosophy 13(1):59–111

Krifka M (1995) The semantics and pragmatics of weak and strong polarity items. Linguistic Analysis 25:209–257

Lahiri U (1998) Focus and negative polarity in Hindi. Natural Language Semantics 6:57–123

Landman F (1998) Plurals and maximalization. In: Rothstein S (ed) Events and grammar, Kluwer, Dordrecht, pp 237–271

Lassiter D (2011) Measurement and modality. PhD thesis, NYU

Lee YS, Horn LR (1994) *Any* as indefinite + *even*, Manuscript, Yale University

Linebarger M (1980) The Grammar of Negative Polarity. PhD thesis, MIT

Linebarger MC (1987) Negative polarity and grammatical representation. Linguistics and Philosophy 10(3):325–387

Nakanishi K (2012) The scope of even and quantifier raising. Natural Language Semantics 20(2):115–136

Nicolae A (2013) Any questions? Polarity as a window into the structure of questions. PhD thesis, Harvard University

Rizzi L (1990) Relativized Minimality. MIT Press, Cambridge, MA

Rooth M (1985) Association with focus. PhD thesis, University of Massachusetts, Amherst

Rooth M (1992) A theory of focus interpretation. Natural Language Semantics 1(1):75–116

van Rooy R (2003) Negative polarity items in questions: strength as relevance. Journal of Semantics 20(3):239–274

Rothschild D (2006) Non-monotonic NPI-licensing, definite descriptions, and grammaticalized implicatures. In: Gibson M, Howell J (eds) Proceedings of SALT 16, CLC Publications, Ithaca, NY, pp 228–240

Schlenker P (2012) Maximize Presupposition and Gricean Reasoning. Natural Language Semantics 20(4):391–429

Schmerling S (1971) A note on negative polarity. Research on Language & Social Interaction 4(1):200–206

Schwager M (2005) Interpreting imperatives. PhD thesis, University of Frankfurt/Main

Schwarz B (2000) Notes on *even*, Manunscript, University of Stuttgart

Stalnaker R (1970) Pragmatics. Synthese 22:272–289

Villalta E (2008) Mood and gradability: an investigation of the subjunctive mood in Spanish. Linguistics and Philosophy 31(4):467–522

Wagner M (2006) Association by movement: evidence from NPI-licensing. Natural Language Semantics 14:297—324

Wilkinson K (1996) The scope of *even*. Natural Language Semantics 4(3):193–215
Yalcin S (2011) Bayesian expressivism. Proceedings of the Aristotelian Society 112(2):123–160.
Yalcin S (2012) Context probabilism. In: Aloni M (ed) Lecture Notes in Computer Science 7218, Springer, Berlin, pp 12–21