1. The licensing question

It is widely believed that negative polarity items (NPIs) are licensed in contexts that invert the normal direction of inference or that license downward entailments (Fauconnier 1975; Ladusaw 1980). NPIs like any and ever can thus occur in the scope of sentential negation, as in (1a), in the restriction of universal quantifiers, as in (1b), and in the scope of negative propositions like without, as in (1c). As the pairs in (2)-(4) show, these expressions all license downward inferences from a superset to a subset case:

(1) a. *I do*(n’t) see anything strange.
    b. *Every*/some student(s) with any sense asked for help.
    c. Nancy crossed the border without/*with any problem.

(2) a. Monica didn’t walk home. |=
    b. Monica didn’t walk home quickly.

(3) a. Every dog was barking. |=
    b. Every brown dog was barking.

(4) a. Nancy traveled without her sons. |=
    b. Nancy traveled without her youngest son.

The generalization that NPIs are licensed in downward entailing (DE) contexts is a major result of the research on negation and quantifiers. This, however, does not mean that it is free of empirical limitations. One that will remain largely undiscussed in this paper is that not all NPIs are licensed in the same kind of DE contexts (e.g. van der Wouden 1997; Zwarts 1998). Another issue the Ladusaw/Fauconnier generalization is known to face is that not all contexts that license NPIs are strictly speaking DE. As (5) shows, if and only, for instance, license NPIs but, as we can see in (6) and (7), they do not license downward inferences:

(5) a. If anyone asks for more coffee, we’ll get some.
    b. Antonella only ever buys Italian shoes.

(6) a. If this match were struck, it would light.
    b. If this match were struck after having been soaked in water overnight, it would light. (Stalnaker 1968)

(7) a. Sid only eats ORGANIC vegetables.
    b. Sid only eats ORGANIC dandelion greens.
Other expressions that license NPIs but require extra assumptions to qualify as creating DE contexts are adverative predicates like regret and be sorry that (Linebarger 1987; Kadmon and Landman 1993; von Fintel 1999, a.o.). Clearly, (9a) can be true without (9b) being true also, for the first example entails (presupposes) that the speaker bought a bike—and it could have been a mountain bike—but the second example requires that he bought a race bike.

(8) *I'm sorry/regret I mentioned anything.*

(9) a. *I'm sorry/regret I bought a bike before I found out about the bike-sharing program.*
    b. *I'm sorry/regret I bought a race bike before I found out about the bike-sharing program.*

The approximatives hardly and barely may also be included in this list (Horn 2002 and references therein). (11a), for instance, entails that Mary passed the test and that she came close to not doing so. That, however, does not mean that she passed the first part of the test and that she came close to not doing so, which is what (11b) requires for it to be true:

(10) *Andrea hardly/barely ate anything.*

(11) a. *Mary barely passed the test.*
    b. *Mary barely passed the first part of the test.*

And, finally, although few and most also license NPIs in their restrictions, their restrictions actually do not license downward inferences in the general case:

(12) *Few/most people who knew anything about the case were willing to come forward.*

(13) a. *Few linguists work on Hungarian.*
    b. *Few Hungarian linguists work on Hungarian.*

(14) a. *Most Americans have a driver’s license.*
    b. *Most Americans under 16 have a driver’s license.*

Yet, despite problems such as these, the Fauconnier/Ladusaw generalization seems so powerful and appealing that the issues just mentioned are generally set aside, with the idea that they can be explained in a way that does not substantially subtract much from the main observation (cf. e.g. von Fintel’s 1999 “Strawson downward entailment”). But even if we manage to address the troublesome cases, the generalization that NPIs are licensed in DE contexts still leaves us with a puzzling question. Why are some expressions only allowed to occur in contexts with particular logical properties? It is not easy to see what lexical choice should have to do with logical inference. In a series of papers Kadmon and Landman (1993), Krifka (1995), Lahiri (1998), Chierchia (2006) and, in a somewhat different form also Israel (1996) argue that the key to the puzzle is to
be found in the meaning of the NPIs themselves. Appealing and elegant though these proposals are, in what follows we argue that they do not fully explain the distribution of NPIs. We argue that while the distributional restrictions on NPIs are indeed related to their meanings, ultimately, the distributional restrictions on NPIs cannot be derived from those meanings. Whether a particular, semantically predisposed element is an NPI is at the end of the day a matter of chance. In the absence of a semantic property that fully determines their distribution, we conclude that what tells NPIs apart from non-NPIs is but a simple, semantically meaningless feature. We then explore several ramifications of our feature theory of NPIs. Our characterization of NPIs in terms of an NPI-feature is shown to tie in well with the historical development of NPIs, including that of ‘even’ NPIs, the NPIs that would seem most amenable to a semantic account. We also discuss how the NPI feature is licensed, relying in this respect heavily on Sánchez Valencia (1991), see also Dowty (1994), Ludlow (2002). We, however, highlight the fact that NPI licensing is only sensitive to what we call Local Polarity and not to Global Polarity. We conclude with some interesting distributional predictions of our feature-based account.1

2. A semantic answer to the licensing question
2.1. The basic idea

Abstracting away from important details, the accounts that aim to derive the distribution of NPIs from their semantics subscribe in some form or another to the following claim:

(15) Semantic answer to the licensing puzzle:
NPIs are bottom scalar elements that associate with/contain an operator that requires them to be the semantically stronger than their scalar alternatives.

Bottom scalar elements occupy the bottom rung on a Horn scale (Horn 1972), e.g. some, or. As is evident from the entailments below, in non-negative contexts some and or represent the weakest of these scalar alternatives and every/all and and the strongest:

(16) a. *Gwen edited all of the articles in the magazine.*
    b. *Gwen edited most of the articles in the magazine.*
    c. *Gwen edited some of the articles in the magazine.*
    
(17) a. *Tom brought wine and beer to the barbecue.*
    b. *Tom brought wine or beer to the barbecue.*

---

1 Earlier versions of our arguments for a feature-based account of NPI licensing were presented at the 2006 Swarthmore Workshop on Negation and Polarity, the 2007 Workshop on Negation and Polarity at the University of Tübingen, and the Parasession on Negation of the 2009 Annual Meeting of the Berkeley Linguistics Society, and were written up in much abbreviated form in Herburger and Mauck (2007, 2009), see also Herburger (2011). The analysis we present in this paper is largely new. We are grateful to the reviewers for their helpful comments and for encouraging us to provide a more detailed account. We would also like to thank Ruth Kramer for helpful discussion of the syntactic issues.
In DE contexts, in contrast, the entailments of course go in the opposite direction and every/all/and now function as the weakest of the alternatives and some (any)/or the strongest:

(18)  
a. It is not true that Gwen edited some (any) of the articles. |=
b. It is not true that Gwen edited most of the articles. |=
c. It is not true that Gwen edited all of the articles.

(19)  
a. It is not true that Tom brought wine or beer to the barbecue. |=
b. It is not true that Tom brought wine and beer to the barbecue.

If, as (15) claims, NPIs correspond to the weakest elements of a Horn scale and come with an operator requiring them to be semantically stronger than any of their scalar alternatives it follows that NPIs are restricted to DE contexts because it is precisely in these contexts where the normally weakest are the strongest. The licensing puzzle seems solved, and elegantly so. As will be discussed next, what requires those low scalar elements that are NPIs to be stronger than their higher ranked alternatives varies somewhat on the different proposals. While the issues discussed in section 3 can be appreciated without knowing the details (and a reader already familiar or not interested in the details can directly skip to that section), we thought it appropriate to describe the specifics of the various accounts.

2.2. The various ‘even’ accounts

The earliest proponents of what can be considered a semantic answer to the licensing puzzle summarized in (15), Kadmon and Landman (1993), claim, first, that NPIs have wider quantificational domains than their non-NPI counterparts and, second, that this widening has to lead to strengthening. Any potatoes is held to differ from some potatoes in quantifying also over potatoes that might normally be considered irrelevant, e.g. old potatoes, very small potatoes, etc. The following dialogue is said to support this (p. 360):

(20)  
YOU: Will there be French fries tonight?
ME: No, I don’t have potatoes.
YOU: Maybe you have just a couple of potatoes that I could take and fry in my room?
ME: Sorry, I don’t have ANY potatoes.

Since the restrictions of existential quantifiers are upward entailing contexts, in non-negative environments a smaller domain results in a semantically stronger claim than an alternative claim involving a larger quantificational domain. Kadmon and Landman’s analysis exploits the fact that in DE contexts the opposite holds; by claiming that NPIs have a wider than usual domain and imposing a requirement that widening must lead to strengthening they predict that NPIs are restricted to negative contexts.

Krifka’s (1995) proposal also builds on the assumption that NPIs like any are low scalar expressions, meaning that any has the semantics of an existential quantifier. For him the
difference between NPIs and their non-NPI counterparts is not how wide their respective domains are, but rather that NPIs (but not non-NPIs) trigger more specific alternatives. For the NPI *anything* these alternatives would be specific things like a pencil, a book, etc. The second part of his proposal is the assumption that NPIs like *anything* appear in the scope of an operator (Scal.Assert) that generates the implicature that no semantically stronger (i.e. more specific) alternative holds. As a result, *Mary saw anything* is ruled out as a contradiction between the meaning of the sentence (‘Mary saw something’) and the implicatures generated by Scal.Assert and the specific alternatives that are posited, namely ‘Mary didn't see a pencil’, ‘Mary didn't see a book’, etc. When the NPI appears in a negative context, no such contradiction ensues and the NPI is acceptable. We are glossing over Krifka’s proposal for strong NPIs like ANY and *at all*, for which he posits a scalar operator with a semantics similar to that of even. In relevant respects this part of the account foreshadows Lahiri’s (1998) analysis.

The beauty of Hindi—and Lahiri’s (1998) analysis of it—is that no silent operator needs to be posited, nor is it necessary to claim that NPIs induce widening or more specific alternatives. This is so because the NPIs in question consist of a focused low scalar element (e.g. *kuch*), which denotes a minimum quantity, say one, and the particle *bhii*, which Lahiri (1998) analyses as a wide scope ‘even’ that generates two conventional implicatures, namely an existential one, saying that some alternative to the focused element (two, three, etc.) makes the sentence true, and a scalar one, saying that the actual amount described in the sentence (one) is the least likely to make the sentence true (Kartunnen and Peters 1979). In positive contexts the semantics of the low scalar *kuch* etc. and the conventional implicatures due to *bhii* are in conflict. If two people came, one person did, so it cannot be that ‘one’ or ‘some’ is the less likely and hence semantically stronger than its alternative ‘two’. If, on the other hand, the NPI is embedded in a DE context, the demands of the wide scope ‘even’ that associates with ‘one’ do not result in contradiction, because now ‘one’ does indeed represent the semantically strongest and hence least likely of the alternatives.²

² Lahiri’s description resembles that in Lee and Horn (1994), who claim that the distribution of *any* is a consequence of it being an indefinite with an incorporated ‘even’, where ‘even’ picks out the least likely of the alternatives. Lee and Horn make use of the fact that the implicatures generated by ‘even’ require the existence of a scale of likelihood associated with the sentence. They then posit that *any* is licensed only where such a scale can be constructed. Whether or not a scale can be constructed can be determined by whether or not even *a single* or even + a superlative can appear in a sentence. Thus, pointing to examples like (i) and (ii), Lee and Horn predict that *any* can occur in a sentence if even *a single* or even + a superlative can occur in the same position in that sentence:

(i)  a. *There is any boy running around in the garden.
    b. *There is even a single boy running around in the garden.

(ii) a. There isn’t any boy running around in the garden.
    b. There isn’t even a single boy running around in the garden.

Lee and Horn also note that many languages, including Hindi, Modern Hebrew, Korean, and Turkish include an instance of even as part of their NPIs.
Finally, Chierchia (2006) aims to give a general account of NPIs, free choice items and scalar expressions in a framework where the pragmatics interfaces dynamically with the syntax. In the spirit of Kadmon and Landman (1993), Chierchia maintains that any denotes an existential quantifier that induces widening in its domain of quantification. And similar to Krifka (1995), Chierchia posits that NPIs come with alternatives, in particular, alternative domains of quantification. In the case of the NPI any (but not a free choice any) these alternative domains involve subsets of the widened domain of quantification of any. He then posits that any bears an uninterpretable feature [+\(\sigma\)], which needs to be checked by a \(\sigma\)-operator. This operator amounts to an instruction to ‘lock in’ the pragmatic implicatures at a given point in the derivation of the sentence. Given the particular alternatives that are associated with the NPI any, Chierchia’s account triggers the presence of an operator \(E_c\), which amounts to a tacit ‘even’. Similar to Krifka (1995)’s Emphatic.Assert, this operator requires that the proposition expressed by the sentence entail all alternative propositions derived with the help of the alternative domains:

\[
E_c(p) = p \land \forall q \in C \left[p \subseteq q\right]
\]

As the alternative domains are smaller than the original, widened one, a proposition involving an element in the original (i.e. wider) domain will only entail the alternative propositions when embedded in a downward entailing context. Thus, \(^*I\ saw\ anyone\) is synonymous with \(I\ saw\ someone\), except that the domain of quantification is wider in the NPI sentence. If the widened domain of quantification consists of \{a,b,c\}, the alternative domains are \{a, b\}, \{a, c\}, \{b,c\}. The \(E_c\) operator then says that the sentence must entail all alternatives generated with the help of the alternative domains triggered by any. But it will fail to do so; if it was \(a\) that the speaker saw, the alternative involving \{b,c\} will be false. This, combined with the action of the \(\sigma\) operator in ‘freezing’ the pragmatics, rules out \(^*I\ saw\ anyone\). By the same token, \(I\ didn’t\ see\ anyone\) is ruled in. The alternatives will be the same as before. Unlike before, however, applying the \(E_c\) operator will not result in contradiction, as all the alternative propositions are entailed by the original proposition; if the speaker did not see any individual from the set \{a, b, c\} then she will not have seen any from the sets \{a, b\}, \{a, c\}, or \{b, c\} either. The widened domains are said to satisfy the requirements of the even-like \(E_c\) operator in downward entailing environments, similar to what we saw in Krifka’s and Lahiri’s accounts.

3. Issues raised by the ‘even’ analyses

Despite their undeniable appeal, the analyses just summarized raise a number of issues, the most obvious perhaps being that they explain distribution of NPIs in terms of felicity but unlicensed NPIs do not so much give rise to infelicity but to ungrammaticality (Giannakidou 2001). Unlike the examples in (22), which are pragmatically bizarre but grammatically well-formed, our initial examples would not be pragmatically bizarre without the licensors but grammatically ill-formed:

\[
\begin{align*}
\text{(22)} & \quad \text{a. } \#\text{Even Dick Cheney wanted to invade Iraq.} \\
& \quad \text{b. } \#\text{Even one person attended the reception.}
\end{align*}
\]
(23) a. *I see anything.
b. *Some student(s) with any sense asked for help.
c. *Nancy crossed the border with any problem.

A more theory-internal problem stems from the assumption that the ‘even’-like operator the advocates of the semantic answer to the licensing puzzle postulate (or exploit) needs to take wide scope over the element responsible for the DE context. If it took narrow scope then for the right truth conditions to follow it could not pick out the least likely or semantically strongest alternative but would need to pick out the most likely or semantically weakest one. While this in itself would not be problematic—it would just mean there are two ‘evens’ (cf. Rooth 1985 a.o.)—it would also mean that the relevant ‘even’ operator is itself an NPI. This would render the analysis circular. Therefore, for the analysis to retain its explanatory force the ‘even’-like operator must necessarily take scope over the element creating the DE context. What speaks against this, however, are instances of NPIs that are licensed across domains that bar scopal movement. It is not clear, for instance, how a tacit ‘even’-like operator lexically associated with anyone could move out of the because clause and above the negation in (24a), or how the two instances of any could move out of the relative clause and above the negation in (24b):³

(24) a. Mary didn’t send the e-mail [because anyone had asked her to] but because she wanted to.
b. They wouldn’t confirm the appointment of a judge [who had voiced any opinion about any controversial matter].

A fourth, quite different concern the semantic answer to the licensing puzzle raises is that English even—and thus perhaps ‘even’ in other languages, including those with overt ‘even’ NPIs—may not really pick out the least likely, semantically strongest alternative but the most informative or noteworthy one, which at times may not be the semantically strongest one. This observation regarding even is due to Kay (1990) and supported by the following example of his:

(25) A: It looks as if Mary is doing well at Consolidated Widget. George [the second vice president] likes her work.
B: That’s nothing. Even Bill [the president] likes her work.

B’s reply indicates that Bill’s approval of Mary’s work is more significant than George’s, even if it is not necessarily less likely (cf. also Herburger 2000). If, however, even really picks out the most noteworthy or contextually significant alternative it becomes difficult

³ This would not be a problem under Chierchia’s account, since for him the ‘even’ operator can be generated anywhere in the syntax as can the σ-operator that checks the [+σ] feature on the NPI. This free generation of σ, however, means that Chierchia cannot exploit what could be an advantage of his proposal: If the feature were not checked, it would actually lead to ungrammaticality rather than the unsuitability that all lexicalist accounts seem to predict.
to derive the distribution of NPIs from the presence of an ‘even’-like operator where the semantics of this operator crucially exploits the inverse relation between likelihood and semantic strength. Of course, one can posit a more abstract operator that is defined to pick out the semantically strongest alternative rather than the pragmatically most noteworthy one. But such a move would seem to undermine the argument from languages like Hindi, where NPIs seem to contain an overt instance of ‘even’.

We may at this point also note that (15) also faces a purely descriptive problem, as also noted Szabolcsi (2004). The proposal is predicated on the assumption that NPIs denote bottom scalar elements. Though this is true in a great many cases, it is not true in all. As discussed in detail in Israel (1996), not all NPIs are low scalar expressions:

(26) \[ \text{La película *(no) fue gran cosa.} \]
    the not ‘I have no idea’

(27) a. \text{He is*(n’t) much of a cook.}
b. \text{It won’t/*(will) be long before they notice.}
c. \text{Writers do*(n’t) accept suggestion much these days.} \text{ (Klima 1964: 283)}

Clearly, \text{gran cosa, much of a, long and much} here are not low scalar expressions but rather relatively high scalar ones. But they function as NPIs nonetheless.

A final reason to remain skeptical of the semantic answer to the licensing puzzle is the existence of pairs like \text{some} and \text{any}. Arguably the (unstressed) NPI \text{any} and the PPI \text{some} have the same meaning (e.g. Krifka 1995, Lahiri 1998). That is, (28b) is the negation of (28a) and (29b) that of (29a):

(28) a. \text{I have some potatoes.}
b. \text{I don’t have any potatoes.}

(29) a. \text{I have potatoes.}
b. \text{I don’t have potatoes.}

The domain widening that can be observed in (Kadmon and Landman 1993) seems to be due to stress on the determiner rather than the semantics of the determiner itself, a point also made in Krifka (1995) and Lahiri (1998):

(30) \text{I don’t have ANY potatoes.}

If, however, \text{some} and the NPI \text{any} are semantically equivalent existential determiners, it cannot be the meaning of \text{any} that is responsible for its distribution as an NPI. Rather, as we have seen, two additional assumptions are needed, namely (i) the assumption that \text{any} and its kin but not \text{some} induce ‘widening’, ‘more specific alternatives’ or ‘higher ranked focus alternatives’, and (ii) the assumption that \text{any} comes with an operator requiring it to
be stronger than its alternatives. Absent strong independent evidence for (i) or (ii), we might as well say that the difference between any and some is that any is an NPI and some is not. Or, put differently, we might say that the NPI any carries an NPI-feature, that must be checked by an element that bears a negation-feature. In contrast with any, some does not bear this feature.

4. The pragmatics of scales and the NPI-feature

Our conclusion that what makes a particular expression an NPI is but a semantically empty NPI-feature raises a number of non-trivial questions, the most obvious of which is perhaps this: If NPI-hood just amounts to carrying a feature, why are so many NPIs low scalar elements? And why are the others relatively high scalar expressions and so much less frequent?

It seems significant that the strength of low scalar items in negative contexts makes them pragmatically useful if what one is after is a strong or emphatic claim (Jespersen 1971; Israel 1996 a.o.). This is clearly true of NPIs that lexically express something of little value (a red cent, etc.), but it also holds true of NPIs that are merely low scalar expressions and synchronically, if not diachronically, lack the colorful content of minimizers, e.g. the NPI any-. We surmise that it is this pragmatic usefulness that leads some low scalar elements to bear an NPI-feature. Which ones do, however, seems to be ultimately up to chance.

Relatively high scalar elements are also pragmatically valuable in DE contexts. But instead of contributing to semantically strong, emphatic, and in some instances hyperbolic claims, they do the exact opposite; they lend themselves to weak assertions, and the attending equivocation, and also understatement (cf. Israel 1996). (31), for example, can be taken to mean that the speaker has some but not much time (weak, equivocal claim) or that she has no time at all (understatement):

(31)  I don’t have much time.

The weak claim arises from a regular scalar implicature due to Quantity, which implicates the negation of stronger alternatives. Since negation reverses the direction of entailment, the stronger alternative to much in (31) is some. (31) therefore implicates that it is not the case that the speaker does not have some time, which means she has some time. It seems reasonable to assume that an understated claim arises when Quantity is suspended and the speaker is taken to not have said everything there was to say. Since without Quantity no scalar implicature is generated in (31), the sentence no longer implicates that the speaker has some time.

The prevalence of bottom scale NPIs as opposed to high scalar NPIs, which has all but obscured the existence of the latter, suggests that we more often put a premium on clarity and emphasis when negation is involved than weakness, vagueness or understatement. In both instances, however, the scalar rung has useful pragmatic consequences in negative
environments. So useful, we argued, that some of the relevant expressions come to be restricted to negative environments.

5. The historic development of NPIs, including ‘even’ NPIs

5.1. Acquisition, change or loss of the NPI-feature

Evidence for the “chanciness” of NPI-hood that we claim to exist comes from the curious fact that a scalar element can be an NPI in one language without being so in another, cf. for instance the English NPI *any* with its merely existential German counterpart *einig-*. It also ties in with the observation that there can be two semantically equivalent elements in one language only one of which is an NPI. As we already saw, English *any* vs. *some* fit this description, as do German *brauchen* vs. *müssen* (‘must’). Finally, if NPI-hood is nothing deep but only a feature, we expect that NPI-hood may not be stable across time.

One way to understand the Jespersen cycle is to say that the NPI-feature sometimes acquires semantic context and turns into a negation-feature. For instance, French *pas* (‘step’) famously developed from an NPI-like companion of *ne* to a negation in its own right in colloquial French. Similarly, English *never*, can be viewed as having an NPI in Old English and Middle English, licensed by *ne*, but now in standard English it functions as a negative expression in its own right. In our terms this means that in both instances the NPI-feature has acquired semantic content and turned into a negation-feature (cf. Herburger 2001).

\[(32)\]  

\[\begin{align*}
He & \text{ nevere yet no vileynye ne seyde In al} \\
\text{his Lif unto no manner wight} \\
\text{his life to n- manner man}
\end{align*}\]  

‘He never said any vile thing in all his life to any man’ (Middle English, 14\textsuperscript{th} century; Chaucer)

The opposite change, namely that of a low scalar NPI becoming a regular indefinite (i.e. that of an NPI losing its NPI-feature feature) does not seem to be as well attested. Jäger (2006, 2007), however, argues that in Old High German, *einig-* functioned as an NPI in the scope of negation, *doubt, without*, in yes-no questions, and in the antecedent of a conditional. In (33) for instance, it appears in the scope of *without*.

\[(33)\]  

\[\begin{align*}
\text{buuzzsan einigan zuuuiun ist dhanne} \\
\text{without any doubt is then}
\end{align*}\]  

\[\text{archennit dhazs…} \\
\text{recognized that}\]

\[\text{\textsuperscript{4}}\] We assume that the Negative Concord that *never* and other currently negative expressions participated in their previous history (and, in some dialects, in current practice) is a phenomenon closely related to NPI licensing. This is a widely held view (cf. e.g. Dowty 1994, Herburger 2001, Szabolcsi 2004 and many others.) It is, however, not a universally held position. Zeijlstra (2004) argues that Negative Concord and NPI licensing are two separate phenomena, see also Penka (2010). Limitations of space prevent us from addressing their arguments here.
‘It is recognized without any doubt that…’ (Isidor VI.5, cited after Jäger 2007)

In its journey from Old High German through Middle High German to Modern German *einig* then must have lost its NPI-feature since now it is used as an existential with free distribution. We take this to indicate that just as NPI-hood can be acquired and transformed it can also be lost. This makes sense if NPI-hood is a semantically inert feature born by some semantically predisposed expression. It makes less sense if NPI-hood is intrinsically linked to the semantics of an expression.

5.2. Analogous development of overt ‘even’ NPIs

Interestingly, the acquisition, change or loss of the NPI feature seems to affect ‘even’ NPIs in much the same way it affects other NPIs. This would be surprising if the distribution of ‘even’ NPIs were determined by the ‘even’ they contain, but is less surprising if ‘even’ NPIs are just like other NPIs and characterized by an NPI-feature.

The change in the NPI-hood of an ‘even’ NPI can be illustrated with data from Modern Hebrew. Like Hindi, Modern Hebrew has morphologically complex NPIs containing an instance of ‘even’ (*af*) that generally require negation to be licensed. This, however, is not the case in elliptical answers, which may suggest that they are already on the path of becoming negative expressions (NEs) in their own right (cf. Herburger 2001):

(34) A:  
who came
‘Who came?’
B:  
even one not
‘No one’

What further lends credence to the suspicion that *af exad* and similar expressions are at this point not only NPIs in Hebrew but also NEs is the existence of expressions where the NPI-feature has acquired semantic content. In particular, in the speech of some children and teenagers *af exad* is starting to be used without negation in preverbal position:

(35) a.  
not bad… even one said that-this bad
‘It’s not bad. No one said that it was bad.’
b.  
…almost even one came with hat
‘…almost no one came with a hat…
[http://www.mynet.co.il/articles/0,7340,L-3715469,00.html](http://www.mynet.co.il/articles/0,7340,L-3715469,00.html)

If the limited distribution of NPIs depended entirely on the internal morphological make-up of the relevant expressions (‘even’ + ‘one’), then, given that this morphological make-
up presumably does not change for individual speakers depending on syntactic environment, the change we described above would be hard to explain. On the other hand, on the present view of NPIs we can say there currently two *af exad* items in the lexicon of Hebrew, one bearing a contentless NPI-feature, meaning ‘anybody’, the other bearing a contentful negation-feature meaning ‘nobody’. In relevant respects *af exad* thus resembles Spanish *nadie*, which has been argued to be currently both an NPI and an NE (Herburger 2001).

Like other NPIs, even ‘even’ NPIs can also lose their NPI-ood. This arguably happened in the historical development of Greek *kanenas*. According to Kiparski and Condoravdi (2006), *kanenas* derives from *kai* (‘and’), *an* (‘even’), and *henas* (‘one’). They report that it was used as an emphatic NPI in early Medieval Greek but ceased to be used as such in later Medieval Greek, where it came to be employed as regular, non-emphatic NPI:

(36)  

\[
\begin{align*}
  kai & \text{ den } eglutösen \text{ kanenas } \\
  \text{and not } & \text{ escaped anyone} \\
\end{align*}
\]

‘And not one of them escaped’ (Makhairas 16.15, cited after Kiparsky and Condoravdi 2006)

What is of special significance here is Kiparski and Condoravdi’s observation that in some dialects of Modern Greek, notably Cretan, *kanenas* is now used as a simple positive indefinite, as shown in (37). It seems then that the ‘even one’ expression *kanenas* in Cretan resembles German *einig* in being a low scalar NPI that has ceased to be an NPI and become a ‘regular’ low scalar expression instead.

(37)  

\[
\begin{align*}
  kianenas & \text{ perastikos da perase } ki \text{ } \dot{\epsilon}\kappa\omicron\rho\omicron\kappa\sigma\epsilon \\
  \text{some} & \text{ passerby here passed-by and cut} \\
  ta & \text{ portokalia} \\
  \text{the oranges} \\
\end{align*}
\]

‘Some passerby must have passed by and cut the oranges’

Finally, the history of some Indo-Aryan languages may offer another example of an ‘even’ NPI reverting to a simple indefinite. Sanskrit formed both indefinites and NPIs by combining a wh-word with a focus particle or a logical connective such as *va* (‘or’). For instance, the combination of *kas* ‘who’ plus *chid* ‘even’ resulted in a word much like English *any*, usable as an NPI as well as a free choice item. The NPI use is illustrated in (38):

(38)  

\[
\begin{align*}
  pravishantam & \text{ na maam kas chid apashyan } \\
  \text{enter-pres.ppl.acc not me-acc who even see-impf.3p} \\
\end{align*}
\]

‘No one (not anyone) saw me entering’ (Mahabharata, Vanaparva, Nala 2,18)

Curiously, in Bhojpuri and Hindi, both descendants of Sanskrit, the neuter form of *kas chid*, *kim chid*, developed into a regular indefinite *kuch* ‘something’ (cf. the *kuch* that appears in the Hindi NPI and free choice element *kuch bhii* discussed by Lahiri 1998):
The development of *kas chid* > *kuch* offers another instance of an ‘even’ NPI losing its sensitivity to DE contexts, i.e. its losing its NPI-feature, and turning to a simple indefinite, which, in turn, may combine with an instance of ‘even’, i.e. *bhii*, to form an ‘even’ NPI.

6. Licensing the NPI feature

6.1. NPI licensing and Local Polarity

Since we concluded that the limited distribution of NPIs is best attributed to a semantically vacuous and, as we just saw, historically unstable feature, we should now address how this feature is licensed. Our task would be more straightforward if NPIs were only licensed under c-command. As, however, Ladusaw’s (1979) discussion of Klima (1964) shows, c-command between licensor and licensee is not the relevant relation as it fails to capture the difference between *every* and *no*; the former indeed licenses NPIs in its syntactic c-command domain (its restriction), but the latter also licenses NPIs in its scope, the syntactic correlate of which the determiner does not c-command. Luckily, in our search for an account with the requisite syntactic properties we can borrow extensively from Sánchez Valencia’s (1991) theory regarding the relation between the monotonicity properties of lexical items and those of the sentences they appear in. The relevance of this theory for NPI licensing is also pointed out in Dowty (1994) and Ludlow (2002), though in a somewhat different way.

Working within categorial grammar, Sánchez Valencia (1991) develops a calculus of the DE-ness or monotonicity of a sentence and its parts. An essential ingredient of his theory is the claim that lexical items are marked for the monotonicity of their arguments. Assuming that determiners denote functions from sets of individuals to sets of sets of individuals, *every*, *some* and *no* carry the lexical features in (40a-c). *Every*, for instance, marks its syntactic sister as downward monotonic and the sister of the QP that dominates it as upward monotonic. Sentential negation bears the feature marking in (40d). (If the determiners are interpreted in situ they denote functions of type <<e,t>, <<e,t>,t>, <e,t>> and receive the corresponding markings in (40a’-c’).)

\[
\begin{align*}
(40) \quad \text{a. every:} & \quad <<e,t>^-, <<e,t>^+, t>> \\
& \quad \text{a’. every':} & \quad <<e,t>, <<e<e,t>^+, <e,t>>> \\
& \quad \text{b. some:} & \quad <<e,t>^+, <<e,t>^+, t>> \\
& \quad \text{b’. some':} & \quad <<e,t>^+, <<e<e,t>^+, <e,t>>> \\
& \quad \text{c. no:} & \quad <<e,t>^-, <<e,t>^-, t>> \\
& \quad \text{c’. no':} & \quad <<e,t>^-, <<e<e,t>^-, <e,t>>> 
\end{align*}
\]
d. **not:** $<<e,t>^-, <e,t>>$

That *every, some, no* and *not* should have the negative features shown in (40) is clearly predictable from their meanings. They carry a minus feature exactly on those arguments for which, speaking now in set-theoretic terms, we can replace one set with a subset without affecting the truth-value. Conversely, we find the plus feature on those arguments where replacing a set with its superset preserves truth. Ludlow (2002) argues that it is possible to dispense with these features entirely. We will discuss this briefly in an appendix.

Since NPI licensing is a syntactic process it is not enough for lexical items to bear polarity features, but these features also have to somehow enter the syntax. In fact, Sánchez Valencia’s (1991) “montonicity marking” describes just how this takes place. We will refer to the process as Local Polarity, to be distinguished from Global Polarity later on. We use “sister” and “mother” in their obvious tree-theoretic sense:

(41) **Local Polarity:**

a. A lexical item whose argument is lexically marked “-” marks its syntactic sister “-”. A lexical item whose argument is lexically marked “+” marks its syntactic sister “+”.

b. All syntactic mothers are assigned “+” except when they themselves are sisters of an expression that assigns “-” to its argument.

As a result of Local Polarity, in (42) the determiner *every* passes on a minus to its sister, the NP. Likewise, *n’t* is responsible for its sister, the VP, ending up with a minus:

(42)
```
      S + t
       |
  QP+ <<e,t>,t>  NEGP+ <e,t>
     /       \
   NEG+     VP- <e,t>
       |       |       |
      V+      QP+ <<e<e,t>,+>,<e,t>>>
        |       |
       Q+ NP-
```

Abelard didn’t catch every’ unicorn

$<<e<e,t>,<e,t>>, <e<e,t>,<e,t>>, <e,t>, <e,t>>$  

This tree, taken from Dowty (1994), assumes that the QP *every unicorn* is interpreted in-situ, in accordance with the theoretical assumptions of categorial grammar. Were it interpreted in a QR-ed position, this would not affect the Local Polarity calculus. The QP *every unicorn* would still contain an NP corresponding to its restriction that bears a minus
and it would still have a sister corresponding to the scope of the QP (the IP the QP adjoins to) that bears a plus.

A third and final step of Sánchez Valencia’s account computes the overall upward or downward monotonicity of a given phrase in a syntactic tree. Let us call this “Global Polarity”:

(43) Global Polarity:
   a. If T is a syntactic tree with root S then a node N has polarity in a tree iff all nodes from N to S are marked “+” or “-”.
   b. A node N has positive Global Polarity iff T has polarity and the number of nodes in the path from S to N that are marked “-” is even. A node N has negative Global Polarity iff T has polarity and the number of nodes in the path from S to N that are marked “-” is odd.

Though the NP node dominating the noun unicorn has negative Local Polarity, its Global Polarity is positive because in the path from NP to S there is an even number of minuses, namely two. By the same token, the V node dominating catch has positive Local Polarity but negative Global Polarity because in the path from S to V we come across only one minus.

Global Polarity is essential for capturing entailment patterns and syllogisms, Sánchez Valencia’s main goal. To give but two simple examples of how it affects entailment, if we replace unicorn in (44a) with an expression denoting a superset of unicorn (mythological creature) as in (44b), the truth of the first guarantees the truth of the latter. Likewise, Global Polarity captures that if (44a) is true, replacing catch with catch and tame preserves truth:

(44) a. Abelard didn’t catch every unicorn.      |=
    b. Abelard didn’t catch every mythological creature.
    c. Abelard didn’t catch and tame every unicorn.

Importantly, however, although piling on an odd number of polarity reversing expressions makes a formerly DE context upward entailing, examples like the following indicate that Global Polarity is of little consequence for NPI licensing. As noted in Ladusaw (1979), NPI licensing requires that the NPI be in a local DE context. It is insensitive to what happens higher up in the sentence. This is effectively also noted in Dowty (1994), where it represents a problem because Dowty “streamlines” Sánchez Valencia’s original proposal by combining Local and Global Polarity. It is also noted in passing in Ludlow (2002):  

---

5 While Global Polarity is not relevant for NPI licensing, it has been argued to be relevant for PPIs in Baker (1970). Baker calls attention to examples where a PPI can appear in the scope of a negation because there is an additional higher negation and suggests that the PPI in examples like the following is OK because the negations cancel each other out.
(45)  

(a) If you don’t say anything, I won’t say anything.
(b) He rarely doesn’t say anything during our monthly meetings.
(c) The Congressman doubted that the former Secretary of State couldn’t remember anything.

Taking the relevance of Local, as opposed to Global Polarity into account and assuming that the minuses discussed above correspond to negation-features, we can now describe the conditions under which NPIs are syntactically licensed:

(46)  

Syntactic answer to the licensing puzzle:

(a) Local Polarity determines the propagation of the lexically marked negation-features in the syntax.
(b) An expression bearing an NPI-feature is licensed iff it is dominated by a phrase bearing a negation-feature.

Since the sister of every, but not that of some, receives a negation-feature through Local Polarity, it follows from (46b) that NPIs are licensed in the restriction of every but not that of some, regardless of whether later Global Polarity later reverses matters for the restriction of every. The same applies to no, but here the sister of the QP also receives a negation-feature from Local Polarity, which means that NPIs dominated by the sister of the QP are also licensed. Finally, no matter what the Global Polarity of VP sister of sentential negation may be, its Local Polarity is negative and its negation-feature licenses an NPI within it.

As we saw when we discussed the semantic answer to the licensing puzzle, NPI licensing is not subject to the same locality constraints other syntactic processes are subject to, cf. (24) above. Barring an intervener of the type discussed in Linebarger (1987), how deeply embedded the NPI is does often not seem to matter. It therefore seems unlikely that the

(i) I don’t think that John didn’t call someone.

Szabolcsi (2004), however, observes that if the negation is in a clause that is higher than the one containing the PPI, no second negation is required. This, she notes, is unexpected on Baker’s cancellation analysis:

(ii) I don’t think that John came up with something.

She further notes that adding a third negation to (i) does not make the PPI ungrammatical, even though the negations no longer cancel each other out:

(iii) I didn’t regret that John didn’t come up with something.

Having rejected Baker’s analysis, Szabolcsi then argues that PPIs are in fact NPIs that need to be licensed twice as soon as there is one licensor present. Like us, Szabolcsi assumes that NPIs carry features. But for her the features are semantically contentful negations that get eliminated through an absorption process. Limitations of space preclude us from discussing the complex details of her analysis here, but it should be clear that in overall spirit, at least, ours resembles hers.
checking of the NPI-feature involves syntactic agreement in terms of feature valuation. This would also be unlikely because the relevant relation between the element bearing the NPI-feature and the element with the negation-feature is domination and not c-command. Moreover, what would correspond to the valued feature (what would be [iNEG]) is higher than the feature that would be seeking the value (what would be [uNEG]), as opposed to the other way around, cf. Chomsky (1999), Pesetsky and Torrego (2004). All of this suggests it would be misleading to call the negation-feature [i-NEG] and the NPI feature [u-NEG] and to assume that the checking of NPI features involves feature valuation. But even if NPI licensing does not have the characteristics of feature valuation, we have presented a number of reasons to think it is a syntactic process nonetheless.  

6.2. Some consequences of feature-based NPI licensing

The observation that Local rather than Global Polarity—and feature licensing rather than inference patterns—matter for NPI licensing may help explain the readiness with which speakers seem to be able to tell whether an NPI is licensed or not. This readiness contrasts strikingly with the difficulties we have in figuring out the entailment properties of sentences once the negative operators start to pile up. Although we have not conducted any experiments, we believe speakers know right away that the NPI anything is licensed in (47). In contrast, computing the meaning of the sentence, which requires calculating Global Polarity, takes a while:

(47)  I don’t think she doubted that the neighbor wouldn’t deny that he had said that he had refused to contribute anything to the new fence.

Since on our analysis the negation-features that check NPI-features are carried by lexical items, we are not stumped by the otherwise puzzling fact that circumlocutions can license downward inferences and not NPIs. As the entailment from (48a) to (48b) shows, like no milk, an amount of milk equivalent to zero is downward monotonic in its scope, but as we can see in (49a) it does not license any.

(48)  a. An amount of milk equivalent to zero/no milk was in the refrigerator.
       b. An amount of milk equivalent to zero/no milk was on the top shelf of the refrigerator.

(49)  a. *An amount of milk equivalent to zero was available in any of the stores.
       b. No milk was available in any of the stores.

The datum an amount of milk equivalent to zero is taken from Szabolcsi 2004 (fn. 19), who attributes it to Paul Postal. As Szabolcsi (2004) notes, its inability to license NPIs

---

6 Zeijlstra (2004) and Penka (2011) develop an analysis of Negative Concord in terms of the checking of [iNEG] and [uNEG] features. They do not extend their analysis to NPI licensing, which for them is an entirely separate phenomenon that they assume is best analyzed in terms of the sort of analyses reviewed in section 2.
represents something of an embarrassment to semantic accounts of NPI licensing. It, however, finds a natural explanation on our view: the lexical item no assigns bears negation-features and assigns them to both its arguments through Local Polarity, the circumlocution \textit{an amount of N equal to zero} does not bear these negation-features and can therefore not assign them to its arguments and license NPI-features contained in them.

Does the relation between semantic DE-ness and NPI-licensing also fail in the other direction? Are there expressions that do not create DE contexts in the strict sense but nonetheless license NPIs and, by hypothesis, bear the negation-feature? At the outset we listed a number of NPI licensors that fail to strictly speaking create DE contexts. Horn (2002) argues that many of them are conjunctive, giving rise to two entailments, an “assertion” entailment and an “assertorically inert” or “backgrounded” entailment. Only the first offers a DE environment but the meaning as whole, namely the conjunction of the two, does not. Expressions of this “secretly conjunctive” sort include only and arguably also every (cf. also Herburger 2000), the approximatives hardly and barely (Horn credits Sevi 1998 with the basic idea), and factive predicates like regret, etc.\textsuperscript{7}

\begin{enumerate}
\item[(50)]
\begin{enumerate}
\item Only Socrates saw it coming.
\item No one other than Socrates saw it coming. \textit{assertion} \item Socrates saw it coming. \textit{backgrounded entailment}
\end{enumerate}
\item[(51)]
\begin{enumerate}
\item Pheidippides barely made it back to Athens. \textit{assertion} \item Pheidippides was close to not making it back to Athens. \item Pheidippides made it back to Athens. \textit{backgrounded entailment}
\end{enumerate}
\item[(52)]
\begin{enumerate}
\item I am surprised/regret you passed. \textit{assertion} \item I expected/wanted you not to pass. \item You passed. \textit{backgrounded entailment}
\end{enumerate}
\item[(53)]
\begin{enumerate}
\item Every bicycle is on sale.\textsuperscript{8} \item \(\forall x \ (\text{Bicycle}(x) \rightarrow \text{On-Sale}(x))\) \textit{assertion} \item \(\exists x \ \text{Bicycle}(x)\) \textit{backgrounded entailment}
\end{enumerate}
\end{enumerate}

What this discussion suggests is that expressions that only create DE contexts in their assertorically active component still license NPIs. We take this as an indication that they are “negative enough” to bear the negation-feature that licenses the NPI-features.

\textsuperscript{7} A different analysis is proposed in von Fintel (1999), where he analyzes the assertorically inert or backgrounded entailment as presuppositions. It would lead us to far afield to discuss the pros and cons of the types of analyses here.

\textsuperscript{8} The backgrounded existential entailment is also present where quantification is over individuals belonging to some intentional realm. This explains why we only assent to the first of the following two sentences (cf. Heim and Kratzer 1998, Herburger 2000):

\begin{enumerate}
\item[(i)]
\begin{enumerate}
\item Every unicorn has a horn.
\item Every unicorn has two horns.
\end{enumerate}
\end{enumerate}
Finally, one may wonder whether *if*, an NPI licensor that is sometimes discussed in connection with *only* and *regret* etc. (cf. von Fintel 1999), should also be accounted for in this “secretly conjunctive” way. We think that despite appearances to the contrary, *if* actually has a downward monotonic semantics; the failure of so-called antecedent strengthening is due to the workings of a tacit *ceteris paribus* condition that is sandwiched between the antecedent and consequent. On this view of conditionals, which is first proposed in Schein (2003) and further developed in Herburger and Mauck (2011), (54) has the logical form paraphrased below. The *ceteris paribus* condition explains why someone who assents to the conditional in (54) will not also automatically believe that if you strike a wet match it will light: normal events of striking a match are ones where the match is dry.

(54) *If you strike a match it will light.*

(55) The possible events where you strike a match are followed by some events where those among them that are *ceteris paribus* events (i.e. where everything relevant was and remained equal or normal) are events where the match lights.

7. Conclusion

We took another look at the claim that NPIs must appear in DE contexts and that the reason they do so is to be found in their internal semantics. In view of several problems we noted, we concluded that the distribution of NPIs is better captured in terms of a meaningless NPI-feature that semantically predisposed expressions can come to bear. We showed how the NPI-feature can change over time and either become semantically contentful and turn into a negation-feature or, alternatively, be lost. We also discussed how the NPI feature is licensed. Building on Sánchez Valencia (1991), we assumed that lexical entries specify the monotonicity properties of their arguments (or parts thereof) and that the resulting polarity markings propagate in the syntax through Local Polarity. We highlighted the fact that Global Polarity, while essential to capturing entailment patterns, has no effect on NPI licensing; for an NPI to be licensed it has to be dominated by a phrase that obtains a negation-feature through Local Polarity. We showed how our account predicts that it may often be harder to tell what the entailment properties of a sentence are than whether an NPI within that sentence is licensed. Having divorced NPI licensing from DE-ness, we were also able to explain why circumlocutions can create DE contexts and at the same time fail to license NPIs. We argued, finally, that we can incorporate into our account in a natural way the view that there are secretly conjunctive lexical items which license NPIs while only being DE in the assertorically active part of their meaning.

One issue that we did not explore in this paper is whether the notion of being “negative enough” can be extended to explain NPI licensing by *few* and *most* or whether the capacity of these determiners to license NPIs should be explained in a different way. We hope to remedy this in future work. Another important issue that we have to leave for
another occasion concerns the various differences between NPIs, for instance those between “weak” NPIs, which can occur in all DE contexts, and “strong” NPIs, which require a licensor that is not only monotone decreasing but also anti-additive (e.g. van der Wouden 1997, Zwarts 1998). On the semantic analyses of NPIs that we reviewed earlier in the paper it is not easy to explain why different NPIs should require different kinds of licensors since all that matters for these analyses is that the NPIs appear in DE contexts, whatever type of DE context it may belong to. In contrast, the differences among NPIs and the licensors they require seems amenable to a feature-based analysis of NPI licensing of the kind we have proposed. On a more detailed version of this kind of analysis, which we hope to develop elsewhere, there may not just be one kind of NPI feature and one kind negation-feature, but there may be in fact several, as has also been proposed in a different context in Szabolcsi (2004). What remains to be explored is what these features are and how they interact with Local and Global Polarity.

Appendix

Following Sánchez Valencia (1991), we have assumed that lexical items mark the polarity of their arguments. The markings are not stipulative, but are clearly predictable from the meanings of the relevant expressions; we obtain a “+” whenever an argument is upward entailing and a “−” when it is DE. But Ludlow (2002) shows that we may be able to dispense with these markings altogether, at least for those markings introduced by quantifiers. These, so the claim, can be directly read off the logical forms (semantic representations), from where they can then be “backwards engineered” into the syntax. We cannot discuss the syntactic part here, as this would lead us too far afield, nor Ludlow’s intriguing treatment of most, but will limit us to briefly illustrating some basic aspects of L*.

Instead of a generalized quantifier analysis, Ludlow proposes a purely first-order analysis of quantifiers which, in order to handle proportional determiners like most, crucially involves quantification over cardinalities. On this theory, the quantifiers over individuals are sub-scripted for cardinality and there are two quantifiers over cardinalities, the existential Σ and the universal Π. When combined with further assumptions, quantification over cardinality results in logical forms where one can tell whether a variable appears in a DE context by simply counting negation operators, which then delivers what we have called Global Polarity. These assumptions are (i) that the only sentential connectives available are negation, disjunction and conjunction (no material conditional), (ii) that disjunction must take scope over negation and conjunction (disjunctive normal form), and (iii) that all quantifiers are fronted (prenex normal form). In L* then ‘Some A is B’, ‘All As are Bs’ and ‘No A is B’ have the following logical forms:

\[(56)\]
\[
\begin{align*}
\text{a. } & \text{Some A is B: } \exists x \geq 1 (A(x) \land B(x)) \\
\text{b. } & \text{All As are Bs: } \forall x \geq 1 (\neg A(x) \lor B(x)) \\
\text{c. } & \text{No As are Bs: } \forall x \geq 1 (\neg A(x) \lor \neg B(x))
\end{align*}
\]
(56a) says there is at least one thing that is both A and B. The fact that neither A(x) nor B(x) is in the scope of negation correlates with the fact that neither the restriction nor the scope of some constitutes a DE environment. (56b) says that all things of which there is at least one are either not A or they are B or both, which rules out that there are things that are A but not B. The fact that A(x) is in the scope of a negation tells us that the restriction of all is a DE context. Finally, (56c) states that all things of which there is at least one are either not A or not B or neither A or B. Here both A and B are directly in the scope of a negation, correlating with the fact that both restriction and scope of no are DE.

References


Horn, Laurence R. 1972  On the semantic properties of logical operators in English. Ph.D. dissertation, UCLA.


Lee, Young-Suk and Laurence Horn 1994 *Any* as indefinite plus *even*. Ms., Yale University.


Sevi, Aldo 1998 A semantics for *almost* and *barely*. MA dissertation, Tel Aviv University.


Zeijlstra, Hedde 2004 Sentential negation and negative concord. Ph.D.
dissertation, University of Amsterdam.