

# THE GRAMMAR OF POLARITY

PRAGMATICS, SENSITIVITY, AND  
THE LOGIC OF SCALES

MICHAEL ISRAEL

*University of Maryland, College Park*

 CAMBRIDGE  
UNIVERSITY PRESS

## 9 *The pragmatics of polarity licensing*

---

Putting together novel expressions is something people do, not grammars.  
(Langacker 1987: 65)

### 9.1 *Affectivity reconsidered*

When Klima first proposed the feature [+Affective] to explain the licensing properties of negatives, interrogatives, and conditionals, it was really just a convenient way of labeling polarity triggers. In creating the label, however, Klima effectively advanced a hypothesis that polarity triggers as a group constitute a natural class in the grammar of English, and by implication in Universal Grammar. The history of polarity studies is in large part an effort to flesh out the intuition behind this hypothesis.

The history having proceeded for over forty years now, it is not too soon to reconsider the hypothesis and the intuition behind it. The assumption that polarity licensing in general can be explained by any single mechanism has already been largely abandoned: theorists who agree on little else – Linebarger (1980; 1987), Progovac (1994), van der Wouden (1997), Giannakidou (2006) – agree in positing distinct licensing mechanisms to account for the distinct sensitivities of different types of polarity items. One might go further, however, and wonder whether the sensitivities of any single polarity item can be adequately captured in terms of a single grammatical feature in the first place. The question is, is polarity licensing in fact a property of some class of linguistic representations (whether syntactic or semantic or of some other sort)? And is it sensible to define affective contexts in terms of necessary and sufficient (or even just necessary) conditions on linguistic structures?

I suggest that the answer to these questions is, if not a categorical *no*, at least a qualified *not really*. The point is not that the notion of an affective context is itself ill-conceived. It all depends on how one conceives it. Trivially, an affective context is one which allows for the felicitous use of a negative polarity item – one in which a polarity item makes a meaningful and coherent

contribution to the use of a well-formed sentence. The question is, what is it about an affective context that gives it these special licensing powers? In this chapter I will argue that affectivity is not a syntactic or a semantic property, and indeed is not a property of sentence types at all; rather, it is an essentially pragmatic phenomenon, and as such is a property of utterances (or of utterance types) and their intended interpretations.

I have proposed that polarity items are scalar rhetorical operators and that polarity sensitivity is in general a sensitivity to scalar semantics. If this is correct, then affective contexts should be distinguished by their scalar properties. In particular, an affective context must:

- i) allow for the construal of an expressed proposition within a coherent scalar model;
- ii) allow for scalar inferences supporting a polarity item's rhetorical force;
- iii) be pragmatically consistent with a polarity item's rhetorical force.

These are conditions on the content and construal of an expressed meaning, and they do not, in themselves, necessarily entail any restrictions on the sorts of linguistic structures which might express such meanings. Affectivity does not depend on a sentence's formal structure but rather on the meaning it is intended to convey and on the rhetorical purposes it is intended to fulfill. It depends not on a sentence's syntactic structure or its semantic representation, but crucially rather on the pragmatic coherence of a sentence's use.

Fauconnier was the first to suggest that polarity sensitivity might involve something other than a structural constraint on grammatical representations. He put it bluntly: "abstract representations play no role in the understanding of the syntactic and semantic facts related to the question of 'polarity'" (Fauconnier 1978: 289). The point, as I understand it, is not that polarity phenomena cannot be formally modeled, but that there is no specifically linguistic level of representation which explains the distribution of polarity items. There is, in effect, no level of linguistic representation – syntactic or semantic – at which the notion of an affective context is defined. Polarity licensing is not a matter of well-formed representations but rather one of coherent conceptualizations: polarity items are licensed in contexts where they make sense and fulfill their conventional expressive functions.

This is not to say that the formal properties of a linguistic context are irrelevant. Logical monotonicity, for example, clearly is a property of linguistically encoded semantic representations, and the monotonicity properties encoded by linguistic constructions clearly affect the inferences one draws from a construction's use. And of course, the syntactic structure of a sentence – for example, whether a polarity item precedes or follows a potential trigger – does

have consequences for the ways in which that sentence can be construed. But while syntactic and semantic properties are important characters in the drama of polarity licensing, the constraints they impose ultimately only set the stage for a story whose hero is the pragmatics of scalar construal.

## 9.2 Scalar construal

Scalar reasoning is a general conceptual ability, and probably a very primitive one. It is not, fundamentally, a linguistic phenomenon. Scalar reasoning depends on an ability to think about a situation in terms of other potential situations, and to draw inferences about potential situations on that basis. While language may exploit this ability and even include forms which depend on its operation (i.e. scalar operators), scalar reasoning itself need not depend on any particular linguistic ability. On the contrary, it appears to be a basic feature of the way people understand their perceptual experiences and organize their conceptual structures, and it is plausible to think that pre-linguistic infants, non-human primates, and perhaps even other animals might share the same abilities. As such, scalar reasoning might be considered an instance of the way humans are able to draw complex inferences based on simple image-schematic structures involving notions like *PATH*, *UP-DOWN*, *CONTAINMENT*, and *INCLUSION* (Johnson 1987; Lakoff 1987; Mandler 1992; Lakoff & Johnson 1999; Bergen & Chang 2005). Or, given the abstractness of scalar reasoning in general, we might view scalar ordering in general as a kind of "superschema" (Grady 2005: 1606) available to structure conceptual domains independently of their particular image (or sensory) contents. Ultimately, whatever the cognitive bases of scalar reasoning, the important point is that this fundamental cognitive ability exists prior to, and independently of, our ability to exploit it in language.

In general, any given situation may be an object of scalar reasoning so long as it can be understood with respect to a contrasting set of similar situations. Understanding a situation in this way does not depend on any objective property of the situation in question, but rather reflects the way a conceptualizer may choose to construe the situation. Scalar reasoning thus depends on a mode of scalar construal and, in this sense, is essentially a non-logical, or pre-logical, conceptual ability: it does not involve the manipulation of objective facts to draw valid inferences, but rather constitutes a form of cognitive pattern completion based on the understanding of a given situation type. Whether or not a given sentence supports a scalar interpretation depends not so much on its logical or referential properties as it does on pragmatic factors which determine

how it will be construed in context. The same expressed proposition might be understood either as contrasting with other propositions within a scalar model, in which case it can be said to receive a SCALAR CONSTRUAL, or simply as expressing information about a particular situation, in which case it can be said to receive a SIMPLE CONSTRUAL.

Consider a situation in which someone utters the sentence *Margaret tasted the jellyfish*. Such an utterance will normally allow a simple construal in which it is taken to express information about a single act of jellyfish tasting on the part of a single individual. But in the right context, and given appropriate background assumptions, the same sentence may also be used to generate scalar inferences. Given that jellyfish, at least in Europe and the United States, is widely considered an exotic and not very appealing sort of food, the assertion here may easily be taken as a comment on Margaret's lack of inhibition when it comes to eating. Such an interpretation probably requires a context in which many foods are available for sampling, and of these, the one which is least likely to be appealing and most likely to be considered repulsive is the jellyfish. In such a context, such a sentence may effectively convey that Margaret was so bold as to taste everything, even the least appealing and most appalling of the various offerings.

It is important to note that at least in its written form the sentence *Margaret tasted the jellyfish* has no overt markers for either a scalar or a simple construal. Its interpretation is, in this respect, underdetermined by its formal structure, and the construal it receives ultimately depends on its fit with the context in which it is used. In fact, several distinct scalar construals are possible, depending on what sorts of situations the sentence is understood as contrasting with. On the interpretation discussed above, jellyfish is understood as contrasting with other, probably more appealing, things one could eat. But one might also understand the same sentence as contrasting Margaret with other people who tasted the jellyfish, or even as contrasting the act of tasting with other things Margaret might have done with the jellyfish, such as merely sniffing it or actually consuming large quantities of it. The first of these would make sense in a situation where Margaret is known to be relatively inhibited in her eating habits, and where the sentence is intended to convey that basically everybody tasted the jellyfish. The second might apply in a situation where Margaret was expected to eat some quantity of jellyfish, and the predicate *tasted* was used to provide a vague indication of how much or how little she actually did eat.

One could, of course, force a scalar construal on any basic sentence. One way would be to insert a scalar focus particle like *even*, either before the intended scalar focus itself (*Margaret*, *tasted*, or *the jellyfish*), or before the

verb, in which case any of the three possibilities can function as the focus. The other would be to indicate the focus prosodically, with a fall-rise intonation on the intended scalar focus. If a particular scalar construal is forced on the sentence, this has significant consequences for the ways in which the sentence can fit into a given context: a given scalar construal has to be compatible with the information structure of the context in which it occurs.

Thus a sentence with subject focus like (1a) can work as an answer to the question *Who tasted the jellyfish?* but not to the question *What did Margaret eat?*; while a sentence with object focus like (1b) can only answer the second of these questions.

- (1) a. Even Margaret tasted the jellyfish.  
b. Margaret tasted even the jellyfish.

Words like *Margaret* and *jellyfish* do not force a scalar construal on their own because they do not inherently contrast with an ordered set of alternatives on a conceptual scale. Polarity items, on the other hand, do. Polarity items are scalar operators profiling a conceptual entity with respect to some set of alternatives ranked on a conceptual scale. As scalar operators, polarity items impose a scalar construal on the interpretation of a sentence, and as such they require a pragmatic context compatible with the scalar construal they impose.

Since scalar construal in general depends precisely on the way a sentence's content is integrated into a larger propositional context, the choice between a scalar and a simple construal is fundamentally pragmatic in nature. However unlike, for example, an implicature, it is not a kind of expressed propositional content itself, but rather a way of accessing expressed content. It is not something that can be said or implicated, but rather a way of saying. As such, though it is fundamentally pragmatic in nature, it can also be grammatically constrained by the presence of a scalar operator. In general, if a sentence is so constrained and yet for one reason or another cannot be accommodated into a larger context, the result is not just pragmatic anomaly, but rather a pragmatically conditioned ungrammaticality.

### 9.3 Logical conditions are not sufficient

The claim that polarity licensing ultimately depends on the availability of a coherent scalar construal and the observation that scalar construal is itself neither a logical nor a structural phenomenon together predict that the constraints on polarity items ultimately cannot be captured in terms of logical or structural conditions. In this section I will consider evidence that the logical properties

of a linguistic context are not sufficient to predict its potential for licensing polarity items. As will be seen, under certain conditions negative polarity items may fail to be licensed in contexts which are downward entailing, and even anti-additive. These licensing failures are systematically linked to pragmatic properties of the sentences in which they appear, and in particular to a sentence's ability to support appropriate scalar inferences for a given polarity item. Basically, there are two important ways that things can go wrong. In many cases, polarity items may fail to be licensed because their use somehow depends on the availability of an incoherent, or otherwise pragmatically anomalous, scalar model. In other cases, licensing may fail because a licenser which logically should work simply does not allow for a scalar construal at all. Both types of licensing failure show that the right sorts of entailment are of no avail if they cannot be construed as applying within a scalar model.

### 9.3.1 *Incoherent scalar models*

Polarity items are scalar operators, and as such they generate scalar inferences. The simple generalization for this section is that polarity items cannot be used felicitously in contexts where the inferences they generate will not make sense.

Although universal quantifiers are uncontroversially downward entailing on their first argument, they do not always manage to license NPIs. Linebarger (1980) and Heim (1984) discuss contrasts like those below.

- (2) a. Every restaurant that charges so much as a dime for iceberg lettuce ought to be closed down.  
b. ??Every restaurant that charges so much as a dime for iceberg lettuce actually has four stars in the handbook.
- (3) a. Anyone who gives a damn about the environment enjoys recycling.  
b. ??Anyone who gives a damn about the environment shops at Ikea.

As Heim notes, the intuitive difference between the (a) and (b) sentences here is that in the (a) sentences, but not the (b) sentences, there is some natural connection between the matrix and relative clauses. As she puts it,

the predicate in (2a) [Heim's 36] is something that applies to restaurants *because* they charge a dime or more for iceberg lettuce ... whereas the predicate in (2b) [Heim's 37] just *happens* to apply to those restaurants without regard to, or even in spite of, what they charge for iceberg lettuce. (Heim 1984: 104–5)

Heim proposes that the NPIs in these examples somehow incorporate the semantics of *even* – in my terms, that they are emphatic scalar operators – and

that as such they only make sense where they can be construed as contrasting with, and more informative than, an ordered set of alternatives. The (b) sentences thus fail here because, given normal background assumptions, they do not supply the necessary set of scalar alternatives that could make the expressed proposition count as emphatic.

Yoshimura (1994) notes a similar phenomenon with respect to NPIs in *before* clauses. Since *before* clauses are downward entailing, and in fact anti-additive (Sánchez Valencia, van der Wouden & Zwarts 1993), a strictly logical view of polarity licensing predicts them to be robust polarity triggers; however, the contrasts below suggest that pragmatic factors have a significant impact on the licensing potential of these contexts.

- (4) a. Miss Prism spilled her wine before she had drunk a drop.  
b. ??Miss Prism poured her wine before she had drunk a drop.
- (5) a. The alarm clock was ringing before I managed to sleep a wink.  
b. ??It was raining before I managed to sleep a wink.
- (6) a. Oscar had been studying linguistics for ten years before he learned a damned thing about pragmatics.  
b. ??Oscar had been fishing many times before he learned a damn thing about pragmatics.

Apparently there is more to polarity licensing than just logical monotonicity. Yoshimura herself accepts the thesis that polarity items are sensitive to monotonicity, and she argues on the basis of these and similar examples that they must be sensitive to a “cognitive constraint” as well. Drawing on Relevance Theory (Sperber & Wilson 1986), Yoshimura claims that NPIs share certain procedural-semantic properties with words like *but* which depend on a contrast between what is said and what might have been expected. A sentence like (4a) is thus licensed, in part, by the fact that one might have expected Miss Prism to have drunk at least a little of her wine before she spilled it; similarly, (4b) is odd because there is normally no expectation that one should drink any before one has poured it.

The point is that the felicitous use of an NPI depends on the availability of some sort of implicit contrast. More specifically, I maintain that NPIs require a context which supports a scalar construal, where their expressed content is naturally understood as contrasting with an ordered set of propositions within a scalar model. Thus the NPIs in the (a) sentences here are acceptable because they form emphatic propositions which implicitly contrast with a range of weaker propositions that could have been asserted. In the (b) sentences, however, the NPIs sound peculiar because normal background assumptions make

it difficult to imagine a situation in which the expressed proposition could be construed as appropriately emphatic. In (6a), for example, the longer one studies linguistics, the more likely one should be to know something about pragmatics: Oscar's minimal knowledge after ten years thus contrasts with what one would have expected him to have learned in that time. But in (6b), there is no particular reason why one should expect Oscar's fishing trips to have had any effect on his knowledge of pragmatics, and so the scalar construal required by the NPI *a damn thing* is difficult to provide without the benefit of some particularly rich context.

The examples in (4) are similarly instructive. There is a natural connection between pouring wine and drinking it: as a rule, given normal social conventions for drinking wine (e.g. don't swig directly from the bottle and don't use a straw), until something is poured, it cannot be drunk. But the relationship here is absolute: waiting longer before pouring something does not, under normal conditions, increase the likelihood that any quantity will be drunk; waiting longer before spilling something, however, may well have this effect. For this reason, the (a) sentence is easily construed within a scalar model, while the (b) sentence is not.

Speakers' intuitions tend to be less robust with more thoroughly negative licensors, but comparable examples may perhaps be constructed for *without*, as in (7), and for sentential negation in (8): again, these examples will not make sense if their minimizers cannot be construed within a coherent scalar model.

- (7) a. Algernon left without saying a word.  
b. ?Algernon napped without saying a word.
- (8) a. Cecily didn't eat a bite of her food.  
b. ??Cecily didn't stare at a bite of her food.

(7a) is appropriately emphatic largely because it is normal when taking one's leave to say at least something; but in (7b) the minimizer *say a word* seems oddly out of place since napping is not something that normally requires talking at all. Similarly, for (8), while there are many activities in which a bite of food might count as a natural minimal unit, staring is not one of them: one can just as easily stare at a whole banquet as at a single bite, and so the minimizer in this context fails to create a particularly emphatic proposition.

The need for a scalar construal is particularly apparent in the *so much as* construction, an emphatic NPI which requires its focus (the nominal or verbal complement of *as*) to denote a very low value within a scalar model. Exactly what counts as sufficiently low is flexible and may vary with context and with

a speaker's rhetorical goals. Still, the examples in (9–10) suggest that while its implementation may be fluid, the requirement itself is fairly strict.

- (9) Rupert wouldn't so much as {look at / talk to / kiss} the cigarette girl.  
(10) ??Rupert wouldn't so much as {run away with / make love to / marry} the cigarette girl.

It is perhaps a trivial matter to accommodate these sorts of scalar constraints within a logical view of polarity. The fact that monotonicity is not always sufficient to license polarity items hardly proves that it is not an important factor. But if the constraints on polarity items systematically exceed the predictions of the Monotonicity Thesis, this might be a sign that the thesis has missed the right generalization. The examples here, in any case, suggest that the essential condition on polarity licensing has more to do with the availability of a well-structured scalar model than with logical monotonicity per se. Where an appropriate scalar model is available, NPIs are licensed; where none can be found, NPIs are unwelcome.

### 9.3.2 *On the need for a scalar construal*

The licensing failures discussed above show that polarity licensing often depends on various sorts of real-world knowledge which may enter into the construction of a well-formed scalar model. But there is also a more fundamental requirement for polarity licensing, which is just that any potential polarity licensor must allow for a scalar construal of the polarity item. That is, it must foster a conceptualization in which the profiled content of the polarity item is construed against a background of scalar alternatives. Most downward entailing contexts naturally allow such construals, but some do not, and when they don't they naturally fail to license polarity items. In this section I will consider one such context.

As is well known, negation in a matrix clause can often license polarity items in a sentential complement. This is true of all neg-raising predicates (e.g. *think, believe, expect*) and of a number of other verbs as well (e.g. *say, realize*). In general, the Monotonicity Thesis predicts the occurrence of NPIs in all these contexts, since they are all demonstrably downward monotonic. The tests in (11), for example, show that the sentential complement of *not realize* is both DE (11a), and in fact anti-additive (11b).

- (11) a. I didn't realize that Monica likes scotch. →  
I didn't realize that Monica likes expensive scotch.  
b. I didn't realize that she smokes cigars or drinks scotch. ↔  
I didn't realize that she smokes cigars and  
I didn't realize that she drinks scotch.

Given these facts, it is not surprising to find that NPIs like *any*, *ever*, and *the least bit* are licensed in this context.

- (12) Ken didn't realize that Monica
- knew any oil executives.
  - had ever been to Haiti.
  - was the least bit interested in glossolalia.

The sentential complement of *not discover* presents an unexpected contrast to what we find with *not realize*. As (13) demonstrates, *not discover* seems to pass the same tests for downward entailment and anti-additivity as *not realize*.

- (13) a. I didn't discover that Monica drinks scotch. →  
I didn't discover that Monica drinks expensive scotch.  
b. I didn't discover that Monica smokes crack or shoots heroin. ↔  
I didn't discover that Monica smokes crack and  
I didn't discover that she shoots heroin.

But surprisingly, NPIs appear not to be licensed in the complement *not discover* – the examples below are especially bad when the NPIs occur with stress.

- (14) He didn't discover that she
- knew {several/\*any} oil executives.
  - had (\*ever) been to Haiti.
  - was (\*the least bit) interested in glossolalia.

Both the verbs *realize* and *discover* are factive in that they presuppose the truth of their complements: (12a) thus presupposes that Monica does know some oil executives, (14b) that she has been to Haiti. The crucial difference seems to be that whereas *realize* can be used as a stative predicate to mean something like 'be aware of,' *discover* only occurs as an achievement verb and so necessarily refers to the single instant in which knowledge is gained. On the relevant reading, when it's used in the past tense, to *not realize* something is effectively to be ignorant of it. The predicate allows a scalar construal because ignorance is an inherently gradable property. *Not realize* licenses *any* here because the sentence as a whole allows for a scalar construal in which the asserted proposition – *K didn't know that M knew even one executive* – contrasts with an ordered set of alternative propositions – *K didn't know that M knew two executives, ... that M knew three executives, ... that M knew four...*, etc. The different propositions within the scalar model make different claims about the depth of Ken's ignorance.

Discovery, however, unlike ignorance, is not a gradient phenomenon. A single discovery does not suggest a set of contrasting discoveries. (14a), like

(12a), presupposes some familiarity with some number of oil executives, and it asserts that this familiarity has not been discovered with even one individual. But here the very nature of discovery prevents a scalar construal. The assertion indicates that a certain type of event, a discovery, has not taken place, but it does not thereby evoke any larger set of possible discoveries against which the assertion is to be evaluated. This fact makes it difficult, if not impossible, to accommodate an NPI here.

#### 9.4 Logical conditions are not necessary

The evidence so far clearly demonstrates that logical monotonicity alone is not sufficient to guarantee polarity licensing. Polarity items, it turns out, are sensitive not just to the logical structure of a linguistic context, but also, and crucially, to pragmatic factors which affect what inferences can be drawn from a sentence on a given occasion of use. This result certainly challenges Ladusaw's strong claim that "the property of being a trigger is completely predictable from the truth conditional meaning of an expression" (1979: 162). On the other hand, Ladusaw himself suggested that downward entailment might be a necessary, but not a sufficient, condition for polarity licensing (1979, 1983: 385). So the question now becomes, can polarity items ever be licensed in contexts which are demonstrably not downward entailing? And if they can, then what role, really, does logical monotonicity have to play in polarity licensing?

In this section I will consider a number of contexts which, at least sometimes, license polarity items, but which are not logically downward entailing. For Linebarger, at least, this result has been taken to show that notions like logical monotonicity are in fact irrelevant to polarity licensing, and to grammar in general. This conclusion may be too uncharitable. Ladusaw, I think, had the right idea but applied it at the wrong level of grammar. Polarity items may not be sensitive to monotonicity per se, but they are sensitive to something very much like monotonicity, that is, scalar inferencing. In the next section I will consider examples which illustrate this sensitivity as it appears in contexts which are not, strictly speaking, downward monotonic.

##### 9.4.1 'Exact' cardinals

Noun phrases involving a precise cardinal quantity (*exactly three puppies, precisely two reasons*) pose a problem for the Monotonicity Thesis. Under the right conditions, an expression of this sort can sometimes license polarity items despite the fact that it is clearly non-monotonic. As the examples below demonstrate, a quantifier like *Exactly 3 N* will not license upward or downward

entailments within its nuclear scope: from the truth of (15a), neither (15b) nor (15c) can safely be inferred.

- (15) a. Exactly 3 professors read a novel last night.  
 b. → Exactly 3 professors read a book last night.  
 c. → Exactly 3 professors read a trashy romance novel last night.

*Exactly 3* is not upward entailing since if just exactly three professors read a novel last night, it still may be that many more were busy reading important scholarly monographs. And *Exactly 3* is not DE since if exactly three professors read novels, it is still possible that they all read different kinds of novels: one might have read a trashy romance, while the other two read trashy mysteries.

Judgments vary on the licensing potential of these sorts of NPs, but I am inclined to agree with Linebarger (1987, 1991) that under the right conditions an 'exact' NP can trigger NPIs: the constructed examples in (16) strike me as fairly unimpeachable.

- (16) a. There are exactly two reasons I would *ever* talk to her again: one is if my life depended on it; the other is if she were to say 'hello' to me.  
 b. Exactly three people have expressed *the slightest* interest in reading my dissertation: you, my mother, and this homeless guy I talk to in the park.  
 c. There are precisely four people in the whole world who would *so much as* consider *lifting a finger* to help that maniac, and one of them is your father.

Linebarger argues that NPIs are licensed in examples like these because, and only because, they convey an appropriate negative implicature. (16a), for example, clearly conveys that beyond the two reasons given, *there are no other reasons* the speaker would *ever* talk to her again. And (16b) similarly conveys that beyond the three people mentioned, *no one else has shown the slightest interest*.

Linebarger's basic intuition in these cases is surely correct. Some notion of implicature must play a crucial role in these examples. At first blush, the examples in (16) can be no more easily explained by pragmatic scalar entailments than they can by logical monotonicity: as the sentences in (17) suggest, *exactly N* is neither scale preserving nor scale reversing, any more than it is upward or downward entailing.

- (17) a. Exactly three professors can solve the hard puzzles.  
 → Exactly three professors can solve the easy puzzles.  
 b. Exactly three professors can solve the easy puzzles.  
 → Exactly three professors can solve the hard puzzles.

In (17a), given three professors who can solve the hard problems, there is no reason to suppose that there isn't an abundance of professors who can

solve the easy puzzles. In (17b), given only three who can solve the easy puzzles, it seems unduly optimistic to suppose that all of them can also solve the harder ones.

Still, it should be clear that the reason examples like those in (16) are well-formed has everything to do with the scalar semantics of exact quantification. The reason an expression like *exactly three N* can license NPIs is that it adds something to what would be expressed by *three N* alone. While *three N* may sometimes be used to express 'at least three N,' *exactly three N* makes explicit, and so indefeasible, the upper-bounding implicature 'no more than three N.' The sense of precision in a word like *exactly* is thus not symmetrical: although in principle it means 'neither more nor less,' in practice the emphasis is often on the 'not more.' And in as much as *exactly three* serves to convey 'no more than three,' it is both downward entailing and scale reversing: if no more than three professors can solve the easy puzzles, then no more than three can solve the hard ones either.

The suggestion here is that sentences like those in (16) license NPIs not because of what they assert or entail, but more generally because of what they convey. And crucially, what they convey here is not just a matter of truth-conditional semantics, but also of a sentence's rhetorical function in context. Unless we allow the 'no more than' reading as a conventional sense of *exactly*, an account based strictly on logical monotonicity cannot explain the licensing in (16). Pragmatic considerations should not affect a monotonicity calculus, but they do affect scalar inferencing. For example, given the (quite natural) assumption that a great many people should be interested in reading a given dissertation, the assertion in (16b) probably conveys more information about how many people will not read this dissertation than about who will. As Linebarger suggests, it is because of this pragmatic potential that NPIs can be licensed in examples like those in (16), and the effect of the NPIs in these examples is itself precisely to emphasize this negative pragmatic potential.

In effect, I am proposing a sort of compromise between Ladusaw and Linebarger. Linebarger is right to point to the importance of implicature in licensing the NPIs in (16); however, her account leaves the scalar nature of the implicature conveniently obscure. Ladusaw is right to point to the importance of inferencing as the crucial mechanism of licensing; however, his account leaves no room for the important role pragmatics often plays in creating the appropriate inferences. The scalar approach to polarity licensing seeks to preserve the insights of both Ladusaw's logic and Linebarger's pragmatics.

## 9.4.2 'Most' and 'few'

The determiners *most* and *few* provide another example of environments which at least sometimes license NPIs but which cannot do so by virtue of their logical properties alone. As Ladusaw points out (1979: 151), *most* is difficult to judge, but the examples in (18) suggest that it is neither upward nor downward entailing on its first argument.

- (18) a. [Most boys who ate an apple] got sick. →  
 b. ← [Most boys who ate fruit] got sick.

(18a) does not entail (18b): it could be that all the boys ate some fruit, and that the apples were poisoned so that most of those who ate an apple got sick, but that really very few boys got sick because most just ate cherimoyas and blackberries and avoided the poisoned apples. This shows that *most* is not upward entailing on its first argument. Similarly, (18b) does not entail (18a): after all, it could be that the cherimoyas were poisoned, but that the apples contained an antidote, so that most of the boys who ate fruit got sick, but those lucky few who ate an apple were all spared. This shows that *most* is not downward entailing on its first argument.

Similar considerations apply to the first argument of the determiner *few*. The examples below suggest that *few*, like *most*, is non-monotonic.

- (19) a. [Few girls who came to the party] rode motorcycles. →  
 b. ← [Few girls who came to the party alone] rode motorcycles.

Again, there are no entailments in either direction. It could be that hardly any girls came to the party on motorcycles, but that all those who did came alone: thus (19a) does not entail (19b). Similarly it could be that few of those who came alone rode motorcycles (say, two or three out of forty), but that most of the girls who came (say, 150 out of 200) came on motorcycles: thus (19b) does not entail (19a).

But while *most* and *few* are both non-monotonic – neither downward nor upward entailing, they do license NPIs (Heim 1984; Jackson 1994; Barker 1995; Israel 1995a). The examples below suggest that many NPIs, both weak and strong, can appear in an NP quantified by *most* or *few*. (20a) comes from Barker (1995: 117); (20b) was heard by Larry Horn in a public service announcement from the NYC Department of Health.

- (20) a. Most children with *any* sense steal candy.  
 b. Most people who have *ever* smoked have already quit, and you can too.  
 c. Most people who have *ever* studied semantics understand the difference between sense and reference.  
 d. Most people who *would lift a finger* to help Bill now are either very foolish or very well-paid.

- (21) a. Few children with *any* sense play frisbee on freeways.  
 b. Few people who have *ever* really thought about it actually believe in the Tooth Fairy.  
 c. Few people with *the least bit* of human feeling can fail to be moved by the heroic achievements of our brave men and women.

While these determiners are logically non-monotonic, in the right context and under the right circumstances, they do allow the limited downward entailments needed to license a polarity item (Heim 1984: 102–4). Thus, for example, *any* is licensed in (20a) and (21a) because these sentences implicate that most children with a lot of sense will steal candy as well and that few children with a lot of sense will play frisbee on the highway either. In other words, despite their logical inadequacy, *most* and *few* can license polarity items because they can trigger scalar inferences.

The examples below demonstrate that both *most* and *few* will, in the right context, reverse entailments in a scalar model. In (22), given the familiar conceptual scale with puzzles ranked in terms of their difficulty, both forms license pragmatic inferences from propositions involving lower-ranked easy puzzles to propositions involving higher-ranked, harder puzzles.

- (22) a. [Most students who could solve the easy puzzles] got a prize.  
 → [Most students who could solve the hard puzzles] got a prize.  
 b. [Few students who could solve the easy puzzles] had trouble on the exam.  
 → [Few who could solve the hard puzzles] had trouble on the exam.

The inference in (22a) is natural because, if prizes are being awarded for solving easy puzzles, one naturally assumes they will also be awarded for solving harder puzzles. Similarly, in (22b), if an ability to solve the easy puzzles is normally enough for students to get through the exam without difficulty, then presumably an ability to solve the harder ones will only make the exam that much more trouble-free. So, given the right context and the right scalar model, the restriction of a quantifier like *most* or *few* does form a scale-reversing context and so licenses the inferences needed to license an NPI.

Of course, the availability of these inferences depends on the context and crucially on the content of the main clause predicate, that is, the quantifiers' nuclear scope. Switching the predicates in these examples switches the inferences they support.

- (23) a. [Most students who could solve the easy puzzles] had trouble on the exam.  
 ← [Most students who could solve the hard puzzles] had trouble ...  
 b. [Few students who could solve the easy puzzles] got a prize.  
 ← [Few students who could solve the hard puzzles] got a prize.



In (23) the inferences now run from hard puzzles to easy ones. In (23a), if an ability to solve the hard problems does not prevent one from having trouble on the exam, then presumably the mere ability to solve easy ones will not help either. In (23b), if students who solve hard puzzles normally do not receive a prize, then presumably students who just solve easy ones will not be awarded either. Thus, the same scalar logic which makes *most* scale reversing in (22a) also makes *few* scale preserving in (23b); and the same scalar logic which makes *few* scale reversing in (22b) also makes *most* scale preserving in (23a).

The inferential properties of these forms thus depend on the ways they are used in context, and, more precisely, on the ways they are construed with respect to a scalar model. Switching the predicates in (22) and (23) switches the scalar models: both models license inferences about students based on their problem-solving abilities, but in one the conceptual scale of puzzles supports inferences about who will get a prize, and in the other it supports inferences about who will have trouble on an exam.

Thus these forms do license polarity items by virtue of their inferential properties, but these are not logical properties of the forms themselves, nor even of the sentences which contain them. They are rather a pragmatic property of the way these forms can be used.

#### 9.4.3 'After' effects

My final example of a polarity trigger whose *affective* powers depend crucially on pragmatic assumptions comes from the occasional use of NPIs with subordinating expressions like *long after* or *hours after*. The following examples, adapted from Linebarger (1987: 370–1), illustrate *after*'s licensing potential.

- (24)
- a. He kept writing novels long after he had *any* hope they might be published.
  - b. The mad general kept issuing commands hours after there was *anyone* left to obey them.
  - c. She kept trying out for the team long after she *had a snow ball's chance in hell of ever* making it.
  - d. He kept wanting to call her up long after he had *the slightest* interest in actually talking to her.

Linebarger (1987, 1991) uses examples like these to support her contention that non-negative polarity licensing depends on the availability of a negative implicature. In each case the situation profiled by the main clause is presented as persisting despite the fact that (i.e. *long after*) some normal or expected condition for its continuation has ceased to hold: (24a) conveys that novel writing

persisted when *there was no hope* of publication; (24b) portrays a general issuing orders with *no one to obey them*; (24c) indicates continued efforts when there is no longer any hope for success; and (24d) concerns a man who still wants to call someone even when he no longer wants to speak with her.

As Linebarger notes, expressions like (*long after*) cannot be considered downward entailing. In fact, the evidence suggests that they are normally actually *upward* entailing. The examples below show that *after* licenses inferences from specific instances (in this case, a specific book) to the general case, and not the other way around.

- (25)
- a. Marcia solved the mystery (*long after*) she read about it in a book. →
  - b. ← Marcia solved the mystery (*long after*) she read about it in *A Natural History of Negation*.

The inference from (25a) to (25b) fails because Marcia might have solved the mystery after reading about it in a book other than Horn's classic tome: this demonstrates that (*long after*) is not a downward entailing operator. The inference from (25b) to (25a), on the other hand, seems generally valid: if she solved the mystery after she read about it in *A Natural History*, then she necessarily solved it after she read about it in a book. This suggests that (*long after*) is in fact upward entailing.

Again, as with *exactly*, Linebarger's basic contention is unimpeachable: the licensing effects in (24) depend on the availability of an appropriate implicature and cannot be attributed to the logical structure of these examples. But what Linebarger fails to appreciate (or at least to acknowledge) is that the implicatures in these cases crucially exploit a scalar logic. Each of these examples depends on a background assumption according to which two types of situations are understood in terms of linked scales. The implicature in (24a), for example, depends on a default assumption that one's interest in writing novels might depend in part on a belief that one's novels could be published. Given this assumption, the contribution of *any* is to emphasize that the novel writing persisted even when the normal motivation of possible publication had been reduced to a minimum. And crucially the use of *any* in an example like this triggers scalar inferences that the novel writing would also persist if there was any greater amount of hope. So the context does succeed in supporting the scalar semantics of the NPI. Similarly, in (24d), there is a normal assumption that the more one wants to talk to someone, the more likely one will be to want to call them. In this context, the emphatic NPI *the slightest* helps to highlight the perversity of the situation by emphasizing the total dissociation of the subject's desire to call from any normal motivating desire to talk. Again, *the slightest* is

licensed because the context does support the inference that if the subject had any greater desire to talk, he would also still want to call.

It seems to me that it cannot be a coincidence that a form like *after* licenses NPIs only in contexts which have this sort of rich scalar structure. Linebarger is correct in pointing to the importance of implicature in these examples. The conclusion, however, is not that inferential properties like logical monotonicity are irrelevant to polarity licensing (*pace* Linebarger 1991 and Giannakidou 1998, 2006), but rather that the relevant inferencing cannot be stated exclusively in terms of a sentence's logical semantic structure.

### 9.5 Rhetorical coherence

Polarity licensing, I maintain, is fundamentally a pragmatic phenomenon, and depends more on the ways a sentence can be construed in context than on its strictly logical properties. Logical monotonicity alone cannot license a polarity item where the appropriate scalar inferences cannot be drawn, and where they can be drawn it is not necessary. It is worth considering why this should be.

The Scalar Model holds that polarity items are scalar operators, and that their conventional semantic properties cause their sensitivities. These properties are themselves fundamentally pragmatic, at least in the broad sense that they concern an expression's subjective construal as well as its objective content. Q-value depends on the way an expression's profiled content is construed against an ordered set of alternatives, and i-value depends on the sorts of rhetorical effect an expression can contribute to an utterance. Polarity items are sensitive because only in certain contexts can they felicitously express both of these scalar semantic-pragmatic properties.

The expression of these features often depends on factors other than just the scalar properties of a licensing context. I-value is basically an index of speaker affect, and as such may be constrained by considerations of politeness or interpersonal rhetoric generally. Emphasis and attenuation, in other words, are not just diacritics which segregate polarity items into scale-reversing and scale-preserving contexts. They are genuine properties of an item's expressive force, with real consequences for its distribution in different expressive contexts. Consequently, polarity items should be sensitive not just to the pragmatics of scalar inferencing, but to the pragmatics of social interaction in general.

Consider, for example, the behavior of polarity items in conditional constructions like those below, where a conditional clause serves a sort of quasi-performative function (Austin 1956; Geis & Lycan 1993).

- (26) a. If you're hungry, there's some pizza in the kitchen.  
 b. If you don't mind my saying so, you look absolutely gorgeous.  
 c. If you've got a minute, could you take a look at this?

Sweetser calls such constructions *speech act conditionals* since "the performance of the speech act represented in the apodosis is conditional on the fulfillment of the state described in the protasis" (1990: 118, emphasis in original). This construction poses a minor problem for any strictly logical approach to polarity licensing, if only because it is difficult to gauge the logical properties of what is essentially a pragmatic construction. While content conditionals clearly do license (at least limited) downward inferences, the inferential properties of speech act conditionals are harder to establish. Thus, while the content conditional in (27) clearly licenses the downward inference from the genus *car* to the species *Lamborghini*, the speech act conditional in (28) does not seem to work the same way.

- (27) a. If you have a car, you can get there on time. →  
 b. If you have a Lamborghini, you can get there on time.  
 (28) a. If you need a car, I'll gladly sell you my old Mazda. ↗  
 b. If you need a Lamborghini, I'll gladly sell you my old Mazda.

In (28) the speaker conditions her offer on the hearer's needs, but just because the hearer's needing a car might justify the speaker's offer to sell her Mazda, it does not follow that the hearer's needing a Lamborghini would also justify such an offer.

Logically, speech act conditionals are very different from their contentful cousins, since in a speech act conditional the truth of the protasis does not count as a sufficient condition for the truth of the consequent. Indeed, in examples like (26a) the sentence as a whole actually entails the apodosis – that is, the pizza must be in the kitchen whether the addressee is hungry or not (Austin 1956: 210; Horn 2000c). But speech act conditionals often do support the sorts of scalar reasoning found with other conditionals, licensing inferences from low to high scalar values. Basically, if something holding to a small degree is enough to justify the performance of a speech act, then, all things being equal, that same thing holding to a greater degree will justify it as well. Thus in (26c), for example, if your "having a minute" is enough to warrant a request for some of your time, then presumably your having even more time would also warrant the request.

Given this sort of logic, speech act conditionals should license NPIs, and indeed they do, at least sometimes. In (29), while the liberal NPIs *anything* and *at all* are fine in clauses conditioning an offer of food, the more emphatic *the least bit* is awkward.

- (29)
- a. If you want anything to eat, there are plenty of eggs.
  - b. If you're at all hungry, there are plenty of eggs.
  - c. ??If you're the least bit hungry, there are plenty of eggs.

Considering just its minimal q-value, one might expect the use of *the least bit* here to emphasize the speaker's sincerity and generosity in offering the eggs: literally, it just seems to say that the offer stands even if the hearer is only minimally hungry. But actually, it sounds oddly manipulative rather than generous.

The problem is that the point of the conditional here is to make the offer without sounding pushy. Speech act conditionals often serve to mitigate the face threat associated with a speech act. In this case, the conditional helps prevent the offer of eggs from being mistaken as a demand that the guest should eat: the offer is conditioned on the addressee's desire to eat, and this gives the addressee a valid option to refuse. But the emphatic nature of *the least bit* closes down that option by insisting on the offer even where it might least be desired. The emphatic NPI thus makes what would be a gracious offer into a peremptory demand.

The examples in (30) illustrate much the same sort of effect. Here the conditional serves to mitigate the face threat associated with an invitation by making it easier to decline it. The choice between *some* and *any* in (30a) may reflect the speaker's optimism that the addressee might say 'yes,' but the more emphatically negative constructions with *any ... at all* and *even a second* in (30b–c) undermine the rhetorical purpose of the speech act conditional, which is to give the hearer options. Instead of making it easier for the hearer to say "no," these more emphatic NPIs suggest that the invitation should be accepted even if he has only the smallest amount of time.

- (30)
- a. Assuming you have {Ø/some/any} time tonight, would you like to see a movie?
  - b. ?Assuming you have any time at all tonight, would you like to see a movie?
  - c. \*Assuming you have even a second to spare, would you like to see a movie?

The overt clash between the insistence of the emphatic NPIs and the deference implied by the conditional invitation creates an awkward effect, and as the polarity items get more emphatic, the effect gets worse.

The problem here has nothing to do with the scalar inferential properties of the conditional construction itself, which does in fact provide the scalar inferences needed to license an NPI. The indirect invitation is predicated on the condition that the hearer should have enough free time, and in general if

having a little time suffices to justify the invitation, then having more time should suffice as well. The conditional does make the high i-value NPIs *any* and *at all* appropriately emphatic. The problem is that the emphatic rhetoric itself conflicts with the deferential pragmatics of the conditional usage as a whole here.

In general, speech act conditionals resist emphatic NPIs. This is particularly true where the premise serves to mitigate the potential face threat associated with an illocutionary act. My judgments in (31) – variations on an example from Quirk *et al.* (1985) – suggest that as NPIs get more emphatic, they undermine the deferential nature of the conditional as a whole and so lead to a decline in acceptability.

- (31)
- a. We're getting married, if it's of any interest to you.
  - b. ?We're getting married, if you're at all interested.
  - c. \*We're getting married, if you're the least bit interested.
  - d. \*\*We're getting married, if you give a flying fuck.

The same effect is evident in (32), and is, if anything, more striking since the NPIs here fail to be licensed by a negation inside the conditional. Note that the semi-NPI *mind* works in this context because it is part of a stereotyped speech act conditional formula, and it fits in the formula precisely because its understating meaning is consistent with the deferential pragmatics of the speech act conditional.

- (32)
- a. If you don't mind my saying so, you should dump that jerk.
  - b. ??If you don't at all mind my saying so, you should dump that jerk.
  - c. \*If you don't in the least mind my saying so, you should dump that jerk.

In these examples, of course, the sentence as a whole can hardly be construed as deferential: the advice given in the apodosis is neither delicately phrased nor likely to be welcomed. Still, the speech act conditional here at least pays lip service to the idea that the speaker will withhold her advice if it might cause the hearer any distress. Once again, the emphatic nature of the NPIs undermines this show of consideration, and again licensing fails because the rhetorical force of the polarity items clashes with the rhetoric of the construction in which they appear.

Emphatic NPIs are not entirely barred from speech act conditionals. They are welcome, in fact, so long as they do not clash with the pragmatic purpose of the conditional itself. Sometimes, for example, a conditional might be used to establish that a given speech act is relevant to the hearer's interest. The example in (33a), from Quirk *et al.* (1985: 1092), is a case in point: the conditional here does not serve to mitigate a face-threatening speech act;

on the contrary, it helps to show why the information expressed in the main clause should be welcome. Emphatic NPIs are more than welcome here, where their rhetorical strength effectively reinforces the generosity of the offer.

- (33) a. I'll be in my office all day, in case you have any problems at all.  
 b. If you have any problems at all, just give me a call.  
 c. If you have even the slightest problem, just give me a call.  
 d. If anybody so much as lifts a finger to interfere with your work here, just give me a call.

The distribution of NPIs in speech act conditionals shows that polarity sensitivity is not just a matter of scalar reasoning *tout court*, but crucially depends on the expression of a polarity item's rhetorical affect. Scalar reasoning is critical because it enables this affective expression, but it is the rhetorical affect itself that determines a polarity item's distribution.

## 9.6 Affectivity reclaimed

The study of polarity sensitivity began in earnest with Klima's notion of affective contexts, and with his hypothesis that there is some "grammatico-semantic" property which unites the class of polarity licensors. Since polarity sensitivity is clearly a grammatical phenomenon, a distributional property of particular linguistic constructions, it has seemed natural to assume that licensing itself is a property of linguistic representations, and that polarity contexts, as a class, are defined by innate principles of grammar which govern the set of possible semantic and syntactic structures. A central purpose of this chapter has been to question this assumption.

The question is not whether polarity contexts form a class of some sort. Obviously they do: they are the class of contexts which license polarity items. However, the evidence reviewed here suggests that polarity contexts are defined not by their formal linguistic properties, but rather by their meaningful pragmatic effects – by the acts of construal in which linguistic forms are used and understood. There is thus no level of grammatical representation, syntactic or semantic, at which there are necessary and sufficient conditions defining the class of polarity contexts – at least not independently of the way such syntactic and semantic structures are deployed and construed in an actual or imagined utterance.

The quality of being a polarity context depends on a combination of morphological, syntactic, semantic, and pragmatic factors, including background

knowledge, contextual assumptions, and speaker intentions, all of which work together to make the use of a polarity item in an utterance actually make sense. But the particular sense polarity items need to be licensed is not some abstract grammatical symbol, but an actual act of scalar inferencing in the construal of what is said.