

Minimizers, Maximizers, and the Rhetoric of Scalar Reasoning

MICHAEL ISRAEL

*Department of English,
University of Maryland, College Park*

Abstract

This paper addresses two basic questions about polarity items: what sorts of meaning can such forms encode and why should such forms exist in the first place. My starting point is the Scalar Model of Polarity (Israel 1996, 1998), which predicts a reliable correlation between a polarity item's sensitivity and its scalar semantic properties: specifically, it predicts that forms denoting a minimal scalar degree may be emphatic negative polarity items (NPIs), while forms denoting maximal degrees can be emphatic positive polarity items. A variety of anomalous polarity items are discussed which flout this prediction, including both emphatic NPIs denoting maximal degrees (e.g. *for all the tea in China, wild horses*) and emphatic PPIs denoting minimal degrees (e.g. *for a pittance, in a jiffy*). The exceptional behavior of these forms is shown to be a direct function of the participant roles they denote, reflecting the fact that different roles license different kinds of scalar inferences depending on how they contribute to the likelihood of an expressed proposition. In addition to establishing a link between thematic structure and the lexical semantics of polarity sensitivity, this result is shown to have important implications for the nature of scalar reasoning generally, and for the role it plays in structuring rhetorical discourse.

0. The Challenge

Polarity items pose a challenge for any simple view of grammar as a functional system. On a functional view one expects that grammar should be, in some sense, useful—that it should make it easier for people to use language. But the constraints on polarity items only seem to make it harder. Polarity items are forms which systematically fail to appear in certain grammatical contexts. They are special precisely because of the ways in which they are, in effect, useless.

Negative polarity items (NPIs), like *the least bit* in (1), can be used in negative sentences, but not in the corresponding positive sentences. Positive polarity items (PPIs), like *rather* in (2), can be used in positive sentences but are awkward, at best, in the corresponding negatives.

- (1) a. Monica isn't the least bit interested in Kosovo.
b. *Monica is the least bit interested in Kosovo.
- (2) a. Bill's new tie is rather attractive.
b. *Bill's new tie isn't rather attractive.

Such restrictions are, on the face of it, perverse. They suggest that a grammar can at times, almost whimsically, bar the use of certain sentences which would otherwise be perfectly meaningful. This is not the sort of thing one expects from a language. Yet languages do include these forms, and often in great abundance. Hoeksema (2000) estimates that there are approximately 500 NPIs in Dutch alone: similar quantities seem likely for German and English at least. And while perhaps not all languages have quite so many polarity items, it seems likely that most, if not all, will have at least some. If grammars really are, in some sense, functionally motivated, it is

mysterious why they should allow the proliferation of such evidently dysfunctional forms.

The perversity of this situation is hardly mitigated by the fact that polarity items can be sensitive to a wide range of contexts beyond negation. As the examples show, the contexts in which NPIs are welcome and PPIs awkward include, among others, the antecedent of a conditional (3), the scope of an interrogative (4), the restriction of a universal quantifier (5), and the standard of a comparative (6).

- (3) a. If you're the least bit nervous we can skip the sky-diving.
b. ??If you're rather nervous we can skip the sky-diving.
- (4) a. Is Monica the least bit worried about the investigation?
b. ??Is Monica rather worried about the investigation?
- (5) a. Anyone who is the least bit interested in verbal periphrasis will want to read this book.
b. ??Anyone who is rather interested in verbal periphrasis will want to read this book.
- (6) a. I'd sooner die than make you the least bit uncomfortable.
b. *I'd sooner die than make you rather uncomfortable.

The problem is not that these contexts are in any sense arbitrary. As the work of Fauconnier (1975a, 1975b, 1976), Ladusaw (1979, 1983), and many others has shown, polarity contexts do share important logical features with respect to the types of inferencing they may support. But these features do not in themselves explain why there should be forms which are sensitive to them. The question is, what is it about polarity items that makes them so sensitive, and where does this sensitivity come from?

The truth is that polarity items are not dysfunctional. In this paper I argue that polarity items exist precisely because they are useful, and that the features which make them so peculiarly sensitive reflect the particular uses which they serve. These uses are essentially rhetorical in nature. Polarity items, I contend, encode a speaker's attitude toward the information she conveys in discourse, and so provide a means of attending to, or otherwise manipulating, the emotional and cognitive responses of the hearer. Polarity items are sensitive to the logical structure of the contexts in which they appear because the rhetorical attitudes they encode crucially depend on the kinds of inferences one might draw from their use. The claim, in essence, is that polarity items are specialized for certain pragmatic functions which in turn make them sensitive to the semantic structures in which they occur, and it is this pragmatically driven sensitivity which ultimately restricts their distributions.

I will offer two lines of evidence in support of this claim. First I will show that the proposed account explains the major patterns of lexicalization characteristic of polarity items as a semantic field. The point here is that by recognizing the pragmatic features which govern polarity sensitivity we may account directly for the range of forms liable to exhibit such sensitivities. The second line of evidence is less direct but may in fact be more important. If the grammar of polarity is driven by general rhetorical principles, then polarity items should offer a privileged view of these principles in action; their effects, however, should also be observable independently, in the rhetorical structure of everyday discourse. In the final sections of this paper, I will thus consider some of the ways these principles may manifest themselves beyond the grammar of polarity sensitivity.

The starting point for this study is the Scalar Model of Polarity (Israel 1996, 1997, 1998), which holds that polarity sensitivity is a sensitivity to scalar inferencing, and that polarity items are themselves a special class of scalar operators. The basic idea is that polarity items are forms which denote an element within a scalar ordering, and which are conventionally used for particular rhetorical effects. It is the

combination of these semantic features which makes polarity items sensitive to polarity. The classic example is the minimizer NPI—forms like *sleep a wink*, *lift a finger* and *the least bit*— in which reference to a minimal scalar degree contributes to the expression of a maximally emphatic proposition. Since the emphatic effect associated with these forms is in fact a conventional aspect of their meaning, they can only be used in contexts where their minimal denotata can somehow count as exceptional or emphatic. This is what makes them polarity sensitive.

The Scalar Model makes clear predictions about what sorts of forms should be polarity sensitive and what sorts of sensitivities they should have. My focus in this paper concerns a small but significant class of polarity items which seem to defy these predictions. These anomalous polarity items encode the same sorts of scalar semantic properties as other polarity items, but their sensitivities are precisely the opposite of what the Scalar Model seems to predict: they act like PPIs when they should be NPIs and like NPIs when they should be PPIs. Thus alongside the minimizers, there exists a parallel class of maximizer NPIs—as in *not for all the tea in China* and *never in a million years*—where emphasis depends on reference to a high scalar degree. The simultaneous existence of maximizing and minimizing expressions serving identical rhetorical functions poses a problem: if sensitivity really is a reflection of scalar semantics, then how is it that forms with apparently opposite scalar semantics can exhibit precisely the same sensitivities?

In what follows I will argue that the problem here is more apparent than real, and that the anomaly of the anomalous polarity items reflects a deeper generalization about the rhetoric of scalar reasoning. As it turns out, the scalar structure of polarity items crucially depends on the roles they play within a larger propositional structure. Different propositional roles may trigger different kinds of scalar inferences depending on the ways they affect the possibility of a proposition being true. It is these roles which determine whether and when a maximal or a minimal amount can add emphasis to a proposition.

The result has implications beyond the grammar of polarity sensitivity. The scalar logic of polarity items is not a peculiarity of grammar, but rather a basic conceptual tool for reasoning about the world, and as such it plays an important role in the rhetorical structure of even the simplest everyday discourse. In this light, polarity sensitivity itself turns out to be just the grammatical reflex of the way a very general conceptual ability, scalar reasoning, influences the structure of discourse.

My argument proceeds as follows. Section 1 introduces the Scalar Model of Polarity and the thesis that polarity items are scalar operators. In section 2 I briefly discuss the relation between the pragmatically-driven Scalar Model account and other theories based on downward entailment or related notions. Section 2 introduces the antithetical, anomalous polarity items which seem to defy the predictions of the scalar model. The synthesis comes in section 3, where both anomalous and canonical polarity items are shown to obey a deeper generalization governing the rhetorical structure of any proposition. Sections 4 and 5 explore the implications of this insight for the logic of scalar reasoning generally and for the rhetorical structure of sentences without polarity items. The conclusion to which I am drawn is that polarity sensitivity itself reflects the interaction of basic cognitive abilities with the pragmatic demands of discourse. In this sense, the grammar of polarity items—that is the conventional linguistic representations which characterize these forms—turns out to involve both semantic (i.e. referential) and pragmatic (i.e. interactional) aspects of meaning which are inextricably bound together.

1. The Scalar Model of Polarity Sensitivity

The Scalar Model of Polarity is a theory of what it means to be a polarity item and what it is that makes polarity items polarity sensitive. The theory draws on a long tradition exploring the relation between grammar, scalar semantics, and rhetoric (e.g. Horn 1972, 1989; Ducrot 1973; Fauconnier 1975a, 1975b, 1976; Anscombe and Ducrot 1983), and is closely allied with work in formal semantics which views

polarity items as sensitive to the inferential properties of logical structures (Ladusaw 1979; van der Wouden 1997; Giannakidou 1998). Crucially, however, the Scalar Model departs from the latter tradition in viewing sensitivity as involving aspects of meaning which are fundamentally pragmatic. Ultimately, I suggest that polarity items are sensitive not so much to the objective truth conditions a sentence encodes, but rather to the rhetorical and subjective ends for which a sentence may be used.

The Scalar Model begins with a simple empirical observation about the sorts of expressions which tend to become polarity items. While polarity sensitivity cannot be predicted on the basis of a form's denotational semantics alone, polarity items do show clear patterns in the sorts of meanings they express. For example, expressions denoting indifference or insignificance (e.g. English *give a damn*, *care to Verb*, *matter*, *be of X account*) frequently become NPIs (cf. Hoeksema 1994); color terms and expressions denoting natural kinds probably never do, at least not unless they function as part of some larger idiom. It would be very surprising to find a language in which the word for a certain kind of fruit or insect could only be used in negative sentences, though such words frequently do function as stereotypical minimal units in idiomatic NPIs like *hurt a fly* or *care a fig*; the point, however, is that where such natural kind terms do appear in polarity items, they do not in fact refer to a natural kind.

More generally, polarity items tend to be associated with certain kinds of pragmatic affect, frequently serving either to intensify or attenuate the rhetorical force of an utterance. For example, among NPIs one finds, on the one hand, emphatic expressions (*lift a finger*, *do a thing*, *say a word*, etc.) which serve to strengthen the force of an expressed or implied negation, and, on the other hand, various sorts of downtoners—forms like *all that*, *so very*, *any too* etc.—whose basic function is to attenuate the force of negation.

These basic and antithetical rhetorical functions, emphasis and attenuation, are characteristic of many, and arguably all polarity items. Interestingly, the expression of these functions systematically interacts with a form's scalar semantics and its particular sensitivity. Emphatic NPIs, as in (7) below, almost always denote (or otherwise make reference to) low scalar values, while attenuating NPIs, like those in (8), consistently refer to high scalar values.

- (7) a. I didn't sleep a wink.
- b. We are not the least bit amused.
- c. She didn't budge an inch.
- (8) a. She didn't sleep much.
- b. He's not all that clever.
- c. This won't take long.

Amongst PPIs one also finds a division between emphatic and attenuating forms, only here, as it turns out, the distribution is reversed: emphatic forms, like those in (9), typically denote high scalar values, and attenuating forms, like those in (10), denote low (or low to mid) scalar values.

- (9) a. She is insanely good-looking, and outrageously cruel.
- b. The test was way too hard. I was utterly depressed.
- c. Boy, are you in a heap of trouble!
- (10) a. She's sorta clever.
- b. I was rather pleased.
- c. They're somewhat worried.

It appears then that there is a simple and consistent correlation between polarity sensitivity and the expression of scalar semantics. Moreover, polarity items seem to

divide rather neatly into four basic types depending on the interaction of two sorts of semantico-pragmatic features: a scalar semantic denotation and an expressive rhetorical force which can be either emphatic, as in (7) and (9), or attenuating, as in (8) and (10). The picture will turn out to be somewhat more complicated when we consider a wider range of examples in section 3, below; for now, however, the four basic types illustrated in figure 1 seem to be a reasonable first pass at a taxonomy of polarity items.

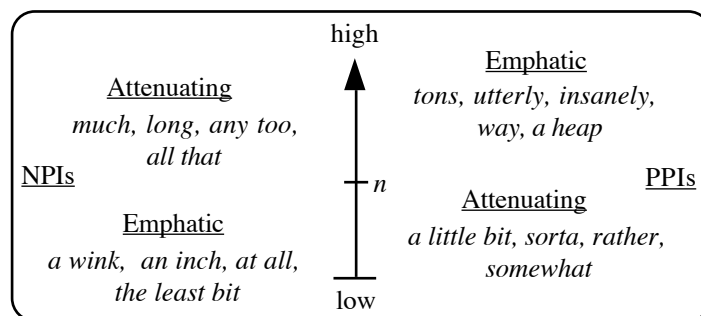


Figure 1: Emphatic and Attenuating Polarity Items

Each of these four types is well-represented in English and in other languages as widespread as Farsi, Japanese, Hindi, French and Dutch (Israel 1998). Even a partial catalogue for English could easily include hundreds of distinct items. And while these forms may vary widely both in their degree of sensitivity and in the precise range of contexts which license or block their occurrence, they are quite consistent in their general semantic range and their basic pragmatic functions.

The robustness of this taxonomy suggests a systematic link between polarity sensitivity and scalar semantics. To be sure, not all polarity items fit this basic pattern quite so transparently as those in the examples above. Still, a great many do, and many which seem less transparently scalar—for example, semi-modals like *need* and *dare*, the punctual use of *until* (as in *we did*(n't) get there until 3*), phasal adverbs like *yet* and *already*—do lend themselves to analysis as instances of or extensions from one or another of the four basic types.¹ At this point, whether or not this taxonomy accounts for all polarity items, the fundamental question is why such a pattern should exist at all. Why should polarity sensitivity be linked to scalar semantics in the first place, and how can a form's scalar semantics explain its sensitivities? As I mentioned above, and as has been amply documented in the literature on polarity items, polarity contexts seem to be defined by the sorts of inferences they license. The trick then is to see how the scalar semantics of polarity items might depend on the inferential properties of different licensing contexts.

Consider an NPI like *the least bit*, a complex scalar adverb which effectively marks a gradable predicate as holding to only the smallest degree. Thus, *the least bit nervous* in (3a) is a predicate denoting the smallest degree of nervousness, and *the least bit uncomfortable* in (6a) is a predicate denoting a minimal degree of discomfort. Strikingly, however, as Fauconnier (1975a, b 1976) demonstrated, the interpretation of these predicates in context is systematically richer than their apparent meanings would suggest. The conditional in (3a), for example, does not require that the addressee be nervous just to a minimal degree for the sky-diving to be abandoned; rather, it suggests that **any** degree of nervousness might justify

¹ For arguments on the scalar nature of phasal adverbs see Michaelis (1993), Israel (1997); for indefinite polarity items like *any* and *some* see Lee & Horn (1994); Haspelmath (1997); Horn (2000); Israel (1995, 1998, 1999); for modal polarity items see Israel (1998); for the punctual use of *until* and its congeners, see Israel (1998), Schwenter (1999), Schwenter & Vasisht (2000).

abandoning the project. Similarly in (6a), one cannot assume that the speaker would happily accept some larger degree of discomfort on the part of the hearer: *the least bit uncomfortable* here effectively means ‘uncomfortable to any degree.’ Apparently then, an NPI like *the least bit* is acceptable in just those contexts where reference to some sort of minimal degree can generate inferences about all other degrees on the same scale. In other words, *the least bit* denotes a minimal scalar degree, but it can only be used in contexts where it licenses inferences over all other scalar degrees and so yields a maximally informative interpretation.

A similar story applies to the PPI *rather*. Like *the least bit*, this form is a degree adverb profiling a limited region on the scale evoked by a gradable predicate. The region profiled by *rather* is hard to pin down, but it clearly lies somewhere above the minimal degree profiled by *the least bit* and somewhere below the high range picked out by an intensifier like *very*. The interesting thing about *rather* is that it seems to be acceptable in just those contexts where it makes an utterance weaker than it would have been with *very*. If I report that my dinner last night was *rather good*, one may reasonably infer that it fell somewhat short of what I might consider *very good*. And it is precisely this sort of inference which makes the b-examples in (3-6) so awkward. That is, it seems difficult to interpret (3b) as meaning that we can skip the sky-diving if you are rather nervous, but we might go ahead with it if you are truly terrified; however, this is precisely the inference which the use of *rather* seems to invite, and indeed, to require. Thus while the NPI *the least bit* requires a context where reference to low scalar values yields inferences about higher scalar values, a PPI like *rather* requires a context in which the inferences go the other way, with reference to high scalar values yielding inferences about lower scalar values.

Apparently then polarity items are forms whose felicity depends on the sorts of inferences one might draw from their use in context. These inferences are scalar in nature and reflect the scalar semantics of the polarity items themselves. Polarity items are “scalar operators” (cf. Fillmore, Kay and O’Connor 1988; Kay 1990): forms whose profiled content is construed against the background of an ordered set of alternatives, and which are thus interpreted within the information structure provided by a scalar model.

A scalar model is a structured set of propositions ordered along one or more parameters in a way that supports inferencing. The model consists of a propositional function with one or more open variables, each ranging over a scale of possible values. The propositional function effectively defines a type of eventuality, and the variables stand for the various ways this eventuality may be realized. A scalar model can be either simple, with one variable and values ordered on a single scale, or complex, with many variables and correspondingly many scales.

Figure (2) presents a simple example with a single variable ranging over a set of puzzles ordered in terms of difficulty. In this case, given the propositional function P , *Norm can solve y*, the model licenses inferences from high values for y to lower values for y . That is, as a rule, Norm’s being able to solve a puzzle y_q of any given difficulty licenses the inference that Norm can also solve any other puzzle y_p lower on the scale of difficulty. Of course, this is not a logical but rather a pragmatic inference: nothing about the structure of the world guarantees it, but all things being equal, the structure of the scale suggests that it should be valid. Fauconnier (1975a) refers to such inferences as “pragmatic entailments” and he calls the scales which support them “pragmatic scales” to distinguish them from Horn’s (1972) quantitative scales based strictly on semantic entailment.

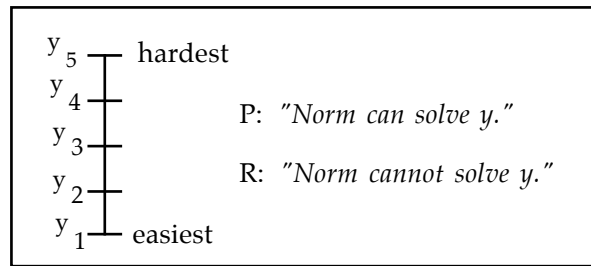


Figure 2: A Scalar Model of Puzzles

Inferencing in a scalar model is defined relative to the propositional function on which it is built. For an affirmative function like P, inferences run from high values to low values on the scale. Normally we assume that if Norm can solve a particularly difficult puzzle he will also succeed with any easier puzzle. With negative propositions, on the other hand, as in R, “*Norm cannot solve y*,” the direction of entailments is reversed, and inferences run from low values on the scale up to higher values. In this case the truth of any proposition $P[y_n]$ low in the model pragmatically entails the truth of all propositions higher than y_n in the model.

Beyond negation, one may draw a broad distinction between those propositional functions which license inferences from high to low values in a scalar model, and those which, like negation, reverse the entailments and license inferences from low values to high values. As Fauconnier (1975a, 1976) notes, this is precisely the distinction needed to define polarity contexts. NPI contexts such as those illustrated in examples (3-6) are always **scale reversing**: they allow inferences from low to high values within a scalar model. PPI contexts, on the other hand, must be **scale preserving**: they support inferences from high values to low values in a scalar model. There are, of course, different ways of formulating this generalization (cf. Ladusaw 1979; Jackson 1994; Zwarts 1996), but the basic idea is by now well-established: polarity licensors—negation, conditionals, questions, universal quantifiers, and others—are united by the sorts of inferences they allow over elements occurring in their scope².

Given the inferential structure of a scalar model, it is easy to see why the two semantic features typical of polarity items—their scalar denotation and expressive rhetorical force—should give rise to polarity sensitivity. As it turns out, both of these features are inherent properties of any proposition in a scalar model. The scalar denotation of a polarity item (or, for that matter, of any scalar construction) determines the position of an expressed proposition within a scalar ordering. Similarly, the rhetorical force of a polarity item is basically just a reflection of the expressed proposition’s informative strength with respect to other propositions in a scalar model.

Assuming that propositions within a scalar model are understood with respect to some scalar norm, we can treat both of these properties as essentially binary features. Relative to the scalar norm, a given form may denote either a **high** or a **low** scalar value. And any expressed proposition will either pragmatically entail the scalar norm, in which case it will count as highly informative (i.e. as saying something which goes beyond some default expectation); or it can itself be entailed by the scalar norm, in which case it will be comparatively uninformative. The informativity of a proposition determines its rhetorical force: highly informative propositions are rhetorically **emphatic**, while weakly informative propositions are rhetorically **attenuating**.

² For more detailed surveys of the inferential properties of polarity contexts see especially Fauconnier (1976), Ladusaw (1980) and Israel (1998). For dissenting opinions on the importance of inferencing in polarity licensing see Linebarger (1987, 1991) and Progovac (1994).

Polarity items are forms which conventionally encode both a particular scalar value and a particular informative strength. As a result, they can only occur where both will be felicitously expressed, and this is precisely what makes them so sensitive. A polarity item with a low scalar value and an emphatic rhetorical force can only occur in those contexts where the expression of its low scalar value will be highly informative. It will therefore be an NPI, limited to contexts where low scalar values entail higher ones. Similarly, a polarity item with a low scalar value and an attenuating rhetorical force will be a PPI: it will only occur in contexts where low scalar values are entailed by higher ones, and where it will thus form appropriately weak propositions.

This, in brief, is the Scalar Model of polarity sensitivity. The theory has three major virtues; it explains the distributions of polarity items directly in terms of their lexical semantics; it provides a taxonomy to account for the major patterns of lexicalization among polarity items; and most importantly, it offers a functional motivation for the existence of polarity items. Polarity items exist, on this account, because the features which make them polarity sensitive are themselves functional elements of lexical meaning. Polarity items' restricted distributions—their apparent uselessness, as it were—are a reflection of the expressive functions which they serve.

The Scalar Model succeeds largely because it recognizes informativity as a natural semantic feature which can be associated directly with a lexical item. Informativity is, in some respects, an unusual sort of semantic feature—really a pragmatic feature, since it does not affect a form's denotation, but rather its conditions of use. Its justification lies in the fact that there are significant expressive and rhetorical effects associated with the informative strength of an expressed proposition. In particular, the use of highly informative propositions can signal a high degree of speaker commitment, or more generally strong speaker involvement in a communicative exchange. Conversely, the use of weakly informative propositions may both help protect a speaker's credibility, and show deference to an addressee by minimizing the demands on his credulity. The basic idea behind informativity as a lexical feature is that these sorts of expressive effects, which arise only in certain contexts of use, may come to be associated directly with a lexical form itself, thus skewing its distribution to just those contexts which favor these effects.

As a theory of what it is that makes polarity items polarity sensitive, the Scalar Model has much in common with recent proposals by Kadmon and Landman (1993), Lee and Horn (1994), and Krifka (1992, 1995), all of which posit features which are at least partially analogous to informativity. In earlier work I have addressed some of the substantive differences among these approaches (Israel 1996, 1998; cf. also Rullmann 1996), and I will not review these here. The most important distinction of the Scalar Model in this group is its breadth of coverage, including both emphatic and attenuating, and both positive and negative polarity items.

The Scalar Model makes clear predictions about what is and is not a possible polarity item, and about the semantic features polarity items must encode. In particular, it predicts that there should be a regular and inviolable association between a form's scalar value, informative strength and polarity sensitivity. In section 3, I will consider a set of examples which seem to challenge this correlation, and which will force us to refine the rather simplistic notion of scalar value introduced above. Before turning to these examples, however, it may help to briefly consider the role of inferencing in a scalar model and how it compares to other notions of logical inference.

2. Scalar Inferencing and Downward Entailment

The idea that polarity contexts are defined in terms of the inferences they license goes back to the work of Fauconnier, but its most influential formulation is due to Ladusaw (1980, 1983), who argued that polarity items are sensitive to the monotonicity properties of logical operators. While the ensuing tradition of monotonicity based analyses (cf. Hoeksema 1983; van der Wouden 1997;

Giannakidou 1998; von Stechow 1999) is sometimes referred to as the Fauconnier-Ladusaw theory, there is in fact a radical difference between Ladusaw's monotonicity based approach and the scalar account first pioneered by Fauconnier and advocated here.

Both types of theory are driven by the insight that polarity contexts are defined in terms of the inferences they support. The theories differ fundamentally, however, in the ways they formulate this insight. For Ladusaw and his successors, the relevant inferences are strictly logical, and the constraints on polarity items are thus taken to hold at a level of logical form representing a sentence's truth-conditional meaning. On the scalar account, the relevant inferences may be either logical or pragmatic, and there is no need to posit logical form as a special level of grammatical representation. The constraints on polarity items are, in effect, just a requirement for coherent conceptual structures. Polarity items require a context which supports scalar inferences consistent with their informativity. But these inferences may themselves be pragmatic in nature, and the licensing of a polarity item may depend as much on pragmatic accommodation as it does on semantic inferencing.

Ladusaw's basic claim is that polarity contexts are created by downward entailing (DE) operators, and that negative polarity items are licensed only when they occur within the scope of such an operator in the logical representation of a sentence's truth conditions. Very roughly, a downward entailing operator is one which allows inferences from general properties to specific instances, from sets to subsets. If, for example, I know that *Beth didn't see a bird on the porch* then I can validly infer that *Beth didn't see a penguin on the porch*. This suggests that sentence negation is a DE operator since it licenses inferences from the general *bird* to the specific *penguin*. This is, of course, a bit of an oversimplification, but it will suffice for present purposes.

Intuitively, the definition of polarity contexts in terms of DE operators is not so different from a definition based on scalar inferencing. Indeed, the set-subset relation used to define downward entailment is itself one of the ordering relations which define a scalar model (cf. Hirschberg 1985). Still, the similarity between the two approaches, though real, is also potentially misleading.

While downward entailment and downward entailing operators are strictly semantic notions, polarity licensing in a monotonicity based theory is fundamentally a structural phenomenon. This is especially clear in Ladusaw's early formulation, where licensing depends on a structural relation between a polarity item and a DE operator at a grammatical level of semantic representation: the constraint is thus formulated as a syntactic constraint on well-formed semantic representations. The theoretical significance of his theory is, in large part, that it helps establish the necessity of such logico-semantic representations in a theory of grammar. This is an extremely significant result, not only because the evidence seems fairly compelling that logical inferences play a role in the grammar of polarity sensitivity, but also because there are few, if any, other grammatical phenomena which seem so clearly to depend on the representation of a sentence's truth conditions.³ On the other hand, the fact that inferencing plays an important role in the grammar of polarity sensitivity does not itself entail that the relevant inferences need to be represented in a sentence's logical form, nor even that they depend on a sentence's literal truth conditional meaning. Indeed, a fundamental claim of the scalar model is that the inferences relevant for polarity licensing regularly go beyond the truth conditions of sentence types alone. If this is correct then the constraints on polarity items cannot be captured within a logical representation of sentence meaning alone, but may depend crucially on the (non-monotonic) ways speaker meaning is constructed in context.

This is not a matter which I hope to resolve in the course of this paper. My purpose here is not to critique the relative merits of monotonicity and scalar

³This point has been brought to my attention independently by David Beaver and Bart Geurts.

accounts, but just to clarify the nature of a debate which might otherwise be ignored. This debate will not be easy to resolve in any case. There are good reasons for thinking that pragmatic aspects of meaning may play an important role in polarity licensing, a point which has been made extensively and repeatedly in the literature (cf. Lakoff 1969; Linebarger 1987, 1991; Heim 1984; Kadmon & Landman 1993; Israel 1996, 1998, in preparation; Horn 2000). But there are always ways of dealing with such cases, ideally by refining the theory (as in Heim 1984; Krifka 1992, 1995; von Stechow 1999), or else, in the spirit of Linebarger, by allowing a role for pragmatics as a secondary licensing mechanism (cf. Giannakidou 1998).

In many ways, the Scalar Model of polarity is more vulnerable than monotonicity based accounts. The scalar account makes two basic predictions which are in principle falsifiable. First, polarity items should be forms whose denotation is construed against a set of scalar alternatives. This seems like a weak requirement, since many semantic domains involve scalar contrasts. But while the theory is permissive (after all, given the diversity of polarity items, it has to be), it does make clear predictions about what sorts of forms can become polarity sensitive—or at least, what sorts of meanings a form can encode as a polarity item. The theory helps explain why, for instance, common nouns and attributive adjectives are generally poor candidates for polarity sensitivity. One typically does not find NPIs with simple meanings like ‘happy’ or ‘red’, or ‘chair’ or ‘oyster.’ The problem is that meanings like these do not saliently contrast with an ordered set of alternatives: *red* contrasts with other color terms, and *oyster* contrasts with other shellfish, but neither of these larger classes are easily construed as forming a scale. When such forms do occur in polarity items they either contribute to a larger idiom which itself involves a scalar construal, or else they take on an idiomatic semantic value themselves. For example, where *red* appears in the idiomatic NPI *a red cent*, it does not denote the distinguishing characteristic of a particular set of coins, but serves as a generic epithet for the class of pennies as a whole. Similarly, in those dialects where *oyster* has functioned as an NPI (as in *it’s not worth an oyster*, cf. Horn 1989: 452) the term does not so much refer to a particular species of mollusk as it does to a generically worthless object. Finally, it is worth noting that the one class of nouns which are regularly prone to polarity sensitivity are measure terms—a class which is, of course, inherently scalar. Here we find in abundance both negatively inclined terms (*jot*, *whit*, *iota*, *tittle*) and positively inclined terms (*tad*, *smidgen*, *mite*, *splash*). What one does not seem to find anywhere are terms which simply denote some natural or artificial kind but which are only meaningful when they occur in negative contexts (or only in positive contexts, and which supplete with synonymous expressions in contexts of the opposite polarity. Given the multitudes of polarity items that have been found, the absence of such forms is striking.

The second basic prediction of the Scalar Model is that polarity sensitivity should reflect the conventional association of scalar predicates with pragmatic affects linked to informativity. If polarity items could be found which either did not encode such pragmatic affects, or for which the combination of scalar meaning and pragmatic force did not predict the form’s sensitivity, the theory would lose much of its appeal. In this case, the lexicalization patterns described in section 1 might be discounted as mere epiphenomena—pragmatically motivated tendencies of a phenomenon which is essentially logical in nature. On the other hand, the Scalar Model receives circumstantial support to the degree that it does successfully predict the lexicalization patterns of polarity items. If polarity items do reliably encode the scalar features predicted by the Scalar Model, then the pragmatics of scalar inferencing provides a natural explanation for both why polarity items should exist and why they should behave the way they do. In general, while monotonicity-based accounts may provide a plausible answer to the latter question, the former (and more fundamental) question is often left a mystery.

In the next section, I will consider a range of polarity items which seem to illustrate the vulnerability of the Scalar Model, but which actually demonstrate its

resilience. Ultimately, these examples will sharpen our understanding of the semantic features which underlie polarity sensitivity, and will further clarify the pragmatic basis of these features in the rhetoric of scalar argumentation.

3. The Anomaly of Inverted Polarity Items

The Scalar Model predicts that there should be four and only four types of polarity items. Although one can never predict whether a given form will encode the relevant scalar features—the association of a semantic feature with a lexical form is always essentially arbitrary—one can at least expect that these features should not interact randomly. Rather, the precise sensitivities of any given polarity item should be a direct function of the scalar features it encodes.

Basically, this means that certain sorts of polarity item should not exist. For example, one should never find an NPI combining a high scalar value with an emphatic rhetorical force—such a combination should only yield PPIs. Similarly, there should be no PPIs combining a low scalar value with an emphatic informative value—such a combination should always create an NPI.

As it turns out, both types of putatively impossible polarity item, or things very much like them, not only exist, but are in fact rather common. Von Bergen and von Bergen (1993: 155-7), for example, point out a variety of ‘maximizing’ NPIs—forms which emphatically strengthen negation precisely by virtue of their *high* quantitative values. Typical instances from English include the underlined examples in (11).

- (11) a. Wild horses could *(-n’t) keep me away.
 b. I would *(-n’t) do it for all the tea in China.
 c. I wouldn’t touch it with a ten-foot pole.

Intuitively, it seems clear that *wild horses*, at least in the idiom in (11a), stands for something like the most irresistible force imaginable⁴. Similarly, in (11b) *all the tea in China* clearly represents an unusually valuable reward, one high on a scale of monetary worth. And in (11c) *a ten-foot pole* (alternatively *a barge pole*) clearly represents an unusually large instrument, one which maximizes the distance between the subject and the thing she touches.

⁴ One might question (as has an anonymous reviewer) whether the idiomatic use of *wild horses* in (11) really qualifies as an NPI rather than just a fixed expression which includes negation: that is, do speakers know the phrase *wild horses* as an idiomatic unit which is restricted to polarity contexts, or do they simply learn the expression *wild horses couldn’t keep NP away* as a sort of frozen saying? Since there is no standard definition, or even rule of thumb, for what should and should not count as a polarity item, this is not an easy distinction to draw. I assume, however, that if a given expression exhibits some minimal flexibility in its usage, and yet appears to be systematically restricted to negative polarity contexts, then it is a negative polarity item.

By this standard, the idiomatic use of *wild horses* clearly qualifies as an NPI. A search of the British National Corpus (BNC) on line (100 million words) turned up 16 instances of the relevant usage involving 4 different verbs: 10 tokens with *drag* (6 involving motion away from a desired state, 4 involving motion toward a distasteful state), 3 tokens with *keep* (*from* or *away*), 2 with *shift*, and 1 with *make ask* (see (i) below). In 14 cases, the NPI is licensed by sentence negation (either by *couldn’t*, *wouldn’t*, or *won’t*), in 1 case by the quasi-negator *scarcely*, and once in the antecedent of a conditional (example (ii)).

- i. Wild horses wouldn’t make her ask what had caused it.
 ii. Even if wild horses dragged her into the police station, nothing could make her talk if she didn’t want to.

The point is that while speakers are at least somewhat flexible in their use of this idiom, they are unyielding in the requirement that it have some sort of polarity reversing license. In this respect *wild horses* is quite similar to the NPI use of *a wink*, which occurs 23 times in the BNC, always as the direct object of *sleep* or *get*, 18 times with overt sentence negation and 5 times with weak or indirect expressions of negation.

Such maximizing NPIs are not peculiar to English. Larrivée (1996) notes parallel constructions from French including *pour tout l'or du monde* 'for all the gold in the world,' *de mémoire d'homme*, roughly 'in living memory,' (cf. also Gaatone 1971: 190), and *de (toute) sa vie* 'in his (whole) life.' Similarly for Dutch Van der Wal (1996) and Hoeksema & Rullmann (2001) point to 'maximum quantity' NPIs such as *voor goud* 'for gold' and *in de verste verte* 'in the farthest distance.'

Parallel to these troublesome maximizing NPIs, we also find a set of equally troubling minimizing PPIs. These are expressions which clearly designate low scalar values and yet produce emphatic propositions in affirmative contexts. The scalar model would seem to make these minimalistic emphatic PPIs just as impossible as the maximalist emphatic NPIs noted above.

- (12) a. Godfrey is (*not) scared of his own shadow.
- b. She would (*not) betray us at the drop of a hat.
- c. You could have knocked me over with a feather.

Clearly, one's own shadow in (12a) is a minimally frightful sort of entity, the dropped hat in (12b) is a stereotypically minimal provocation, and the feather in (12c) is a minimally forceful sort of instrument for knocking something over. And yet just as clearly the use of these expressions in (12) creates a conventionally emphatic sort of utterance.

While inverted polarity items may be less common than their canonical counterparts, they are not particularly rare either, and in certain semantic fields they may actually be the norm. For example, Hoeksema (1997), points out that both English and Dutch feature an open-ended class of inverted NPIs denoting large time spans, as in (13). And as it turns out, there is also a large class of inverted PPIs denoting minimal spans, like those in (14).

- (13) a. We have*(n't) heard from you in a coon's age!
- b. You'll *(never) in a million years guess who I saw last night.
- (14) a. We will (*not) be back in a jiffy.
- b. I would(*n't) marry that woman in a New York minute.

Comparable examples include PPIs like *in a flash*, *in a second*, and *in a heartbeat*, and NPIs like *in days*, *in weeks*, *in years*, *in ages*, and *in a blue moon*.

It gets worse. With polarity items encoding notions of monetary value, both high and low scalar values can be rhetorically emphatic with either polarity. Thus one finds unambiguously emphatic NPIs denoting things of little value (canonical: *a red cent*, *a plugged nickel*, *a thin dime*, *a brass farthing*) and unambiguously emphatic NPIs referring to things of extreme value (inverted: *for all the tea in China*, *for all the money in the world*, *for love or money*, *for the life of me*). Similarly, there are emphatic PPIs referring to things of the greatest value (canonical: *a king's ransom*, *an arm and a leg*) and other, equally emphatic PPIs referring to things of the most negligible value (inverted: *for peanuts*, *for a song*, *for a pittance*). The examples below illustrate this pecuniary paradox of polarity sensitivity.

- (15) a. He won't spend a red cent on your wedding.
- b. She wouldn't kiss him for all the tea in China.
- (16) a. Julio spent a king's ransom on the party.
- b. But he somehow got Madonna to play for peanuts.

These examples, and others like them, show unequivocally that the simple correlation between scalar semantics and polarity sensitivity cannot be as simple as one might have hoped, or as the Scalar Model would seem to have predicted it should be. Maximizing NPIs and minimizing PPIs appear to invert the normal correlations observed among the more canonical polarity items. As figure 3, below, suggests, the

existence of both canonical and inverted polarity items would seem to preclude the possibility of there being any regular correlation between scalar semantics and polarity sensitivity.

On the other hand, all of these apparent counterexamples do share a clearly scalar semantics with their canonical counterparts: that is, their rhetorical strength in use is systematically linked to the scalar values which they express. In effect, these forms incorporate an implicit ‘even’ in their meanings (*I wouldn’t do it for love or money* means ‘not even for love or money’, *Godfrey is afraid of his own shadow* means he is afraid ‘even of his own shadow’): they can only be used where they can be construed as contributing to a relatively informative (i.e. emphatic) proposition within an ordered set of alternatives. The distribution of emphatic inverted polarity items thus seems to obey the same scalar logic which rules the distribution of emphatic canonical polarity items, and as with the canonical items this scalar logic is driven by the pragmatics of informativity.

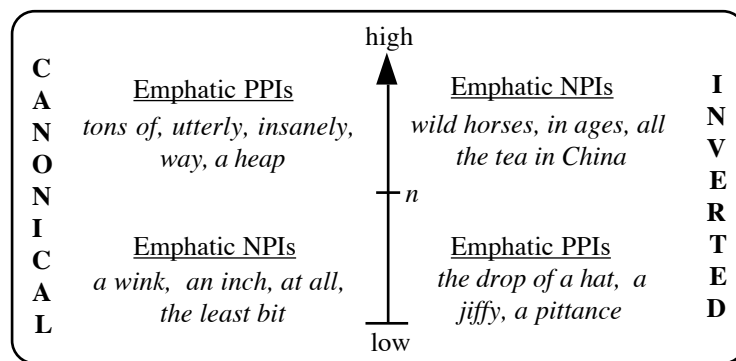


Figure 3: Canonical and Inverted Polarity Items

As with any emphatic polarity item, the inverted forms are only acceptable where their use licenses inferences of the appropriate sort. In (15b), for example, *all the tea in China* marks an extreme value on a scale of incentives, and under negation its use triggers inferences for all values lower on the scale: presumably, if the girl wouldn’t kiss him for all the tea in China, nothing less could tempt her either. Similarly for Godfrey in (12b), if he fears such minimally fearsome things as his own shadow, we may safely infer that he will also be scared of anything more fearsome—that is, in effect, of anything.

The question is, how do these forms, which invert the normal scalar semantics of canonical polarity items, still manage to obey the same scalar logic?

4. The Thematic Logic of Scalar Models

The solution to this mystery is not far to find. The fact is that inverted polarity items are not randomly distributed over the lexicon of polarity sensitivity. Rather, there is a consistent correlation between the role a polarity item plays within a proposition and its status as inverted or canonical.

We may note first that inverted polarity items seem to involve indefinite NPs in a restricted range of semantic and syntactic functions. Prototypical minimizers—forms like *crack a book*, *hurt a fly*, *lift a finger*, *bat an eye*, *sleep a wink*, *budge an inch*, and *breathe a word*—feature indefinite direct objects which measure out the action of the predicate. NPIs of this sort quite consistently display a canonical scalar semantics. Inverted polarity items, on the other hand, tend to feature idiomatic NPs anywhere but direct object position. The *wild horses* idiom, for example, is exceptional in being, to my knowledge, the only English NPI with an idiomatic subject NP. Other inverted items involve indefinites governed by prepositions such as *for* (*for love or money*, *for a song*), *in* (*in a flash*, *in a million*

years), with (*touch with a ten-foot pole, knock down with a feather*) and at (*at the drop of a hat, at a moment's notice*).

Underlying this superficial syntactic distinction there is a deeper semantic generalization. A pattern begins to emerge when one considers the thematic roles typically associated with canonical and inverted polarity items. Canonical polarity items tend to refer, in one way or another, to a patient (*crack a book, hurt a fly*), a theme (*lift a finger, move a muscle, bat an eye*), or more generally some sort of increment (*sleep a wink, drink a drop, budge an inch, breathe a word*). All these forms involve entities which are somehow affected by the action of the verb: they are low in the thematic hierarchy, or, in other terms, near the bottom of the action chain (Langacker 1987).

Inverted polarity items, on the other hand, tend to involve participants at the top of the thematic hierarchy—entities which somehow play a causal, or at least a facilitating role in the realization of an eventuality. The idiomatic use of *wild horses*, for example, denotes a stereotypically irresistible force which might prevent one from the proper exercise of his or her own will. Thematically, the *wild horses* idiom fits into a more general class of inverted polarity items which refer to a stimulus or causal trigger for some eventuality. Other such forms include *at the drop of a hat* (a small event which provokes a significant reaction) and *scared of one's own shadow* (a negligible threat which triggers significant fear). Pushing the generalization a bit, forms like *all the tea in China* and *for a song* also denote a sort of stimulus, referring as they do to potential rewards which might motivate one to act in a certain way. Finally, polarity items involving reference to an instrument (*touch with a ten-foot pole, knock down with a feather*) are always inverted: the use of a bigger or more powerful instrument tends to facilitate the performance of an act.

It appears then that the division between canonical and inverted polarity items reflects a deeper distinction in the ways scalar reasoning applies to different propositional roles. Certain types of participants function effectively as obstacles to the occurrence of an event; others, on the contrary, act as stimuli. A theme or patient, for example, is an entity which must be affected for an event to take place: the bigger it is, the more resistance it offers, the less likely the event will be. An agent or a stimulus, on the other hand, is an entity which must itself be effective for an event to take place: in this case, the bigger the agent, the more powerful it is, the more likely the event will be.

In this light, the paradox of pecuniary polarity items (cf. 15-16, above) simply reflects the fact that the valuables in a commercial exchange are split between two very different sorts of participant roles. As a rule, any participant in a commercial exchange must both give something up and gain something in return: otherwise, the exchange won't happen. We may thus distinguish between the valuables given and the valuables gained. The logic of self-interest treats these two types of valuable very differently. All things being equal, any rationally self-interested participant will strive to give up the smallest amount necessary, and to gain the greatest amount possible. The logic of commercial exchange thus depends on whether a given valuable is understood as a Resource—what one stands to lose—or a Reward—what one stands to gain. The greater the demands on one's Resources, the less likely one will be to engage in an exchange; conversely, the greater the potential Reward, the more likely one will be to go forward with the exchange.

Given this, it seems clear that canonical polarity items—emphatic PPIs encoding large quantities (*an arm and a leg, a king's ransom*) and emphatic NPIs encoding small quantities (*a red cent, a plugged nickel*)—refer to Resources⁵, while inverted polarity items—emphatic PPIs encoding small quantities (*for a song, for peanuts*)

⁵ Resources may appear in (at least) two distinct participant roles, either as expenses (as in *It won't cost you a red cent*) or as possessions (as in *We don't have a red cent*). I am grateful to an anonymous reviewer for helping me recognize this generalization.

and emphatic NPIs encoding large quantities (*for all the money in the world*)—refer to Rewards. Fundamentally then canonical and inverted polarity items do obey the same scalar principles. Emphatic NPIs, whether canonical or inverted, always pick out a class of participants—in this case, big rewards and small expenses—which will facilitate the realization of a given event type. Emphatic PPIs, on the other hand, always denote participants of a sort which militate against the realization of an event—in this case, small rewards and large expenses.

Similar considerations apply to the logic of temporal polarity items like those in (13-14), where emphatic NPIs often denote large time spans (*in days, in weeks, in years, etc.*) and emphatic PPIs often denote minimal spans (*in a flash, in a sec, in a jiffy, etc.*). Of course, there is nothing about the domain of time itself which makes long time spans emphatic in negative contexts and short ones emphatic in positive contexts. As it turns out, the briefest moments can also be emphatic in negative contexts, as in (17), and the longest periods can be emphatic in affirmative contexts, as in (18).

- (17) a. I {won't/??will} be half a minute.
 b. I can *(not) for a second believe she would do that.
- (18) a. It has always been so, time out of mind.
 b. They've been married for {ages /an eternity /donkey years}.

The key to this apparent chaos is, again, the realization that the same type of entity—in this case, a temporal interval—may be associated with very different sorts of roles within a proposition.

In the case of time spans, the crucial difference depends on the aspectual character of an expressed proposition. Basically, whether or not a long time span makes a given eventuality more or less likely depends on the durativity of that eventuality. Punctual events culminate in an instant within a temporal interval: the longer the interval, the more likely it is that the event will actually happen. Durative situations, on the other hand, must hold for every instant of a time span: the more time that passes the more likely it is that the situation will no longer obtain.

So in (17) and (18), where the temporal expressions indicate how long a situation will or will not last, the expression of emphasis involves a canonical scale: brief durations are emphatic under negation; long durations are emphatic in affirmations. Inverted forms like those in (13) and (14), however, invariably designate the bounded interval within which an event takes place. Their inverted scales—with short intervals emphatic in affirmation—simply reflect the logic of punctuality: the shorter the interval the less likely it is to include the moment where a punctual event takes place. Again, the apparent anomaly of these forms turns out to be a regular feature of the roles they play within an expressed proposition.

The question remains, though, why certain propositional roles are associated with canonical scales and others with inverted scales. As I noted above, there are suggestive correlations here with traditional thematic roles like agent, patient and instrument. But thematic structure alone is an unwieldy instrument for sorting out canonical and inverted polarity items. Aside from the fact there is no consensus on the inventory of thematic roles, or even on their status within linguistic theory, it is unclear, at best, how the multiplicity of thematic roles should map onto a binary distinction between inverted and canonical scalar semantics. Put bluntly, the question is what do Agents, Stimuli, Instruments, Rewards and Temporal Intervals all share that distinguishes them from Themes, Patients, Expenses and Durations?

One might think of the distinction broadly as a force dynamic division between “antagonistic” participants (agents, stimuli, etc) which facilitate the realization of an eventuality, and “agonistic” participants (patients, themes, etc) which act against the force of an antagonist to impede the realization of an eventuality (cf. Talmy 1985). The explanation seems appealing with a contrast like the one between the affected fly in *hurt a fly* and the forceful horses in *wild horses*, but other polarity items do

not lend themselves so naturally to a force dynamic analysis. In an expression like *have a clue*, for example, it is hard to see how the clue acts against or impedes the ‘having’ relation. Similar considerations apply to the paper in *be worth the paper it’s written on* and the ghost in *stand a ghost of a chance*. And it seems a stretch to think of temporal intervals like *a jiffy* or *a million years* as forces compelling or impeding the occurrence of an event.

Another possibility might be to appeal to Dowty’s (1991) notion of proto-agent and proto-patient to predict when a polarity item can be inverted. One might thus propose that polarity items are inverted if and only if they occur in argument positions where they have more proto-agent properties than at least one other argument⁶. Such a proposal should make the right predictions for at least those canonical polarity items involving a direct object. It also seems to work for the canonical NPI *be worth the paper it’s written on*: since both arguments of the predicate *be worth* entail just one of Dowty’s proto-agent properties (i.e. neither one is more agentive than the other), the NPI is canonical. Still, it remains unclear, at best, how the proposal can work for temporal polarity items like *a coon’s age* and *in a jiffy*, which do not seem to entail any proto-agent properties. And then, even if one could extend the general framework to handle such cases, it is unclear what one would gain thereby, unless one could also somehow explain what it is about proto-agentivity that makes it invert polarity items.

Ultimately, in order to understand the division between canonical and inverted polarity items we have to understand the roles such forms play within the structure of a scalar model. A scalar model is basically a conceptual tool for thinking about the relations between different possible eventualities. The structure of the model is such that if one knows the status a given eventuality (i.e. whether it does or does not hold), one may automatically infer the status of other, related eventualities within the model. This in fact is the key to the problem of inverted polarity. Elements on any scale within a scalar model are always ranked in the same way, that is in terms of the entailments they yield for a given propositional schema. Those elements which in scale preserving contexts form the propositions with the most entailments are ranked at the top of the scale; those elements which under the same conditions form the propositions with the fewest entailments are ranked at the bottom. The ranking thus does not depend on the objective properties of the scalar elements alone, but is crucially determined by the way these properties interact with a given propositional schema.⁷

Normally, of course, one thinks of scales more concretely as ordered in terms of amounts or degrees. Canonically, these orderings run from lesser to greater amounts. Prototypical scales measuring things like size, weight, or intelligence regularly

⁶ Dowty’s proto-agent properties are (1) volitional involvement, (2) sentience or perception, (3) causing an event or change of state in another participant, (4) movement relative to another participant, (5) independent existence from the event denoted by the verb.

⁷ Loosely speaking, one might say that elements in a scalar model are ranked in terms of the probability that they will satisfy a given propositional schema—that is, in terms of their likely participation in a certain type of situation. The most likely, and therefore least remarkable participants in a given situation always start at the bottom of the scale; the least likely, and most remarkable participants come at the top. Technically, as Paul Kay has pointed out (p.c.), this way of talking about things may well be unworkable: statistical probabilities are very different sorts of creatures from scalar inferences, and in any case it is not clear how one would actually assign probabilities to the distinct elements in a scalar ordering. Strictly speaking, the structure of a scalar model depends on the entailment relations which it assigns to its constituent propositions, and not on the relative probabilities of different possible situations. Still, Kay’s caveat notwithstanding, I suspect that our basic experience of what makes a given situation more or less likely must undergird our understanding of the scalar models we use to reason about that type of situation. For this reason, I may occasionally allow myself the liberty of speaking loosely about the likelihoods with which different elements may participate in a given situation.

conform to this pattern, and the pervasive scalar metaphor, MORE IS UP, whereby an increase in amount is conceptualized as a rise in elevation (cf. Lakoff and Johnson 1980), similarly presupposes a canonical ordering of elements from lesser to greater quantities. But the canonical scale is in fact just a special case (albeit the default case), and in this case, as with all others, the ordering depends on the scale's role within a larger propositional frame. The frame here involves nothing more than the attribution of a scalar property: *x has property to extent-y*. Such scales are always canonical, running from smaller to larger extents. Their logic reflects the fact that if some entity instantiates a property to some high degree, then it must also instantiate that property to all lesser degrees as well. This sort of logic makes these scales useful in reasoning about the possible existence of different entities (cf. Hoeksema & Rullmann 2001). Thus, for example, while things may vary infinitely in weight, everything with any weight will weigh at least a minimal amount, but relatively few things will weigh as much as a ton: so things with a minimal weight will be more likely to exist, and the canonical order for weights will run from light things to heavy things.

The distinction between canonical and inverted polarity items depends on the fact that for every canonical scale there exists a corresponding inverted scale. And which of these two scales is used in any given scalar model depends on the role it plays there. So, again speaking of weights, a propositional frame like *this camel can carry X* has a very different logic from a frame like *X will break this camel's back*. The first requires a canonical scale, the second an inverted one. And the choice in both cases depends on the way the variable affects the possibility of the proposition as a whole being true.

In short, it is an irreducible fact about scalar logic that different roles within a proposition will have different effects on the likelihood of the proposition's being true. Some roles involve entities which may facilitate the realization of a proposition; others involve entities which militate against its realization. Canonical polarity items always involve roles of the latter sort; inverted polarity items always involve roles of the former sort. The scalar logic in both cases is identical, and in both cases it depends strictly on the way propositions within a scalar model are ordered in terms of their entailment relations.

Inverted polarity items do not undermine the scalar model of polarity sensitivity. On the contrary, they confirm it. They do, however, raise questions about the basic structure of scalar models and, more particularly, about the nature of quantitative value. The basic generalization remains that polarity items encode a fixed position within a scalar model, and quantitative value is the feature which expresses this positioning. However, the relevant notion of quantity is not a matter of size or amount *per se*, but rather reflects the particular role which a given referent may play in the realization of a proposition. Loosely speaking, one may think of quantitative value as reflecting the probability that a given element on some scale within a scalar model will yield a true proposition. Scalar models themselves constitute complex presuppositions about the way the world usually works, and the orderings of elements within a model simply reflects our default understanding of how those elements will contribute to the realization of a given situation type.

To conclude, polarity items are defined not just with respect to the contexts which license them, but also in terms of the roles they express in a given propositional function. Different propositional roles may be associated with different scalar orderings. More precisely, the ordering of elements for any given propositional role in a scalar model depends on the way that role affects the possibility of the proposition as a whole being true. This is a non-trivial result and its implications go well beyond the parochial domain of polarity sensitivity. In the final sections of this paper I will consider some of these implications for the rhetoric of scalar reasoning in general.

5. Coordinated Scales in Complex Scalar Models

It should be clear that scalar models are not just an obscure grammatical device for regulating the distribution of polarity items. Rather, they are a basic conceptual tool for reasoning about the world, and they are a pervasive feature of everyday discourse. Although they usually seem too obvious to merit attention, the structures underlying such reasoning can be surprisingly baroque.

As noted above, a scalar model may consist of one or more scales, depending on the number of variables in its propositional frame. Complex scalar models with multiple scales are particularly useful in reasoning about the relation between various elements in a proposition. They also help to illustrate the way the ordering of any given scale depends on its role within the proposition as a whole.

In their analysis of the English *let alone* construction, Fillmore, Kay & O'Connor (1988: 520) discuss the car-wash example in (19), where five distinct scalar parameters are coordinated in a single complex proposition.

- (19) You'd never get a poor man to wash a car for \$2 in bad times, let alone a rich man to wax a truck for \$1 in prosperous times.

The relevant five parameters here are informally represented in a schema like *M would do W on O for R under C*. Each variable in the schema defines a distinct scale, as follows:

- M = individuals, ranked in terms of their financial need.
- W = types of work, ranked in terms of strenuousness.
- O = objects to work on, ranked in terms of size.
- R = rewards for work, ranked in terms of financial value.
- C = general circumstances, ranked in terms of the availability of work.

The rhetorical structure of the sentence and the meaning of the conjunction *let alone* require that all five scales be coordinated in a way that will maximize the contrast between the two conjuncts of *let alone*. In effect, the semantics of *let alone* requires that for each variable, the value in the first conjunct (i.e. *a poor man, washing, a car, \$2, and bad times*) must be ranked lower than its corresponding value in the second conjunct (i.e. *a rich man, waxing, a truck, \$1, and good times*). As with all scalar models, elements on each parameter are ordered in terms of the likelihood that they will participate in a situation of the relevant sort. Since a rich man is less likely to need work than a poor man, elements on M are ordered from the least to the most prosperous people. And since people will work more when they have more to gain, elements on R are ordered with large rewards (\$2) at the bottom and smaller rewards (\$1) higher up.

Two things stand out about this example. The first is its remarkable complexity; the second is the ease with which it is understood. Understanding an example like this depends on the unconscious and automatic coordination of multiple scales in a single scalar model. The fact that this is so easy to do should not make it any less impressive. On the contrary, it shows that the most mundane features of everyday cognition may involve minor miracles of computational complexity.

If anything, the car-wash example makes scalar reasoning seem more exotic than it really is. As it turns out, the same complex processes of scale coordination can be found at work in some of the simplest monoclausal sentences. Consider the examples in (20-21), where the schema *X can solve Y in Z-time* provides the scaffolding for exclamations about the problem-solving abilities of people, the solubility of problems, and the time spans in which an event occurs.

- (20) a. *Stella can calculate an eigen vector in the blink of an eye.*
b. *Even Dim could solve *this puzzle* in a minute.*
c. *With that much time even I could do the Sunday crossword.*

- (21) a. *Dim* can't even count to two in less than an hour.
 b. Even Stella couldn't solve *this puzzle* after a year of trying.
 c. *It all happened faster* than Einstein could say $E=MC^2$.

The logic of these examples again seems obvious. Stella is extremely bright because she solves difficult problems quickly. Dim is extremely dull because it can take him ages to do the simplest thing. But these particular examples instantiate a more general, and rather less obvious formula, the intricate rhetorical structure of which tends to be obscured by the ease with which it is understood.

Each sentence expresses amazement about the remarkable qualities of some topical element (the one in italics) by associating it with stereotypically extreme values on two coordinated scales (the ones underlined). In each case, the expression of wonder depends crucially on the polarity of the sentence as a whole: it would not be so surprising if Stella *could not* solve the hardest problem or if Dim *could* solve the simplest. But the logic of these examples also depends crucially on the proper coordination of the relevant scales in a well-constructed scalar model. In this case the three scales are:

X = Problem-solvers, ordered from most competent to least competent.

Y = Problems, ordered from the least to the most difficult.

Z = Time spans, ordered from longest to shortest.

Elements on each scale are ordered in terms of the likelihood that they will satisfy the propositional schema *X can solve Y in Z*. A difficult problem, for example, ranks high on the scale of problems because it is unlikely to be solved; conversely, a clever problem-solver like Stella ranks low because she is likely to solve any given problem.

In general, in order to emphasize the extreme position of any given element on a scale within the model, one must associate it with elements from the opposite ends of other scales within the model. Specifically, in order to emphasize that some element E is low on a given scale, one asserts that E will satisfy a proposition in which very high values are assigned on other scales; conversely, to emphasize that an element is high on a given scale, one denies that it will satisfy a proposition with very low values assigned on other scales. In other words, an improbable (i.e. high) value for any given scale is improbable to the degree that it is unlikely to satisfy a proposition even with the most probable values assigned for every other scale; and a probable (i.e. low) value is probable to the degree that it is likely to satisfy a proposition even with very improbable values on the other scales.

As often happens, the general rule seems more complicated, and more difficult to understand than its particular instantiations. Scalar reasoning in everyday language seems such an obvious affair—so easy to understand, one scarcely sees it. But as it turns out, scalar reasoning is not obvious at all. It only seems that way because we employ it so effortlessly, so automatically and so often.

It is worth noting that the complexities of scalar reasoning in everyday discourse are precisely the same complexities which underlie the use and distribution of polarity items. The rhetorical structure in examples like (20-21) depends on the analysis of a complex proposition into multiple scalar parameters, each ordered in terms of its effect on the likelihood of the proposition being satisfied. These parameters must then be coordinated and deployed in a linguistic context whose polarity supports the scalar inferences needed to produce an appropriately emphatic utterance.

In utterances like those in (20-21), where the rhetorical purpose is an emphatic expression of wonder, expressions denoting low scalar values are effectively like negative polarity items: they can fulfill their rhetorical purpose only if they occur in an appropriately negative (or, more precisely, scale reversing) context. Similarly, expressions denoting high scalar values function like positive polarity items: they

lend emphasis to an utterance only when they occur in affirmative, scale preserving contexts.

Thus, as the Scalar Model predicts, the distinction between NPIs and PPIs simply reflects the different rhetorical potentials associated with different scalar values. And the distinction between canonical and inverted polarity items reflects the way the ordering of elements depends on the role they play in a scalar model. Canonical scales, like the scale of problems for (20-21), involve elements that make a situation less likely as they increase along a parameter. Inverted scales, like the scales of problem-solvers and temporal intervals, involve elements which make a situation more likely as the increase along a parameter.

The need to modulate one's utterances to achieve particular rhetorical effects is a basic feature of linguistic interaction. The conceptual structure of scalar reasoning turns out to be particularly useful in this respect. By carefully manipulating scalar inferences a speaker can modulate the rhetorical strength of her utterances. The use and distribution of polarity items thus turns out to be the grammatical reflex of a much more general cognitive phenomenon. Polarity items are lexical forms which have been specialized for the expression of scalar rhetoric. The fact that polarity items are so widespread and plentiful, both within English and crosslinguistically, is itself a testament to the use and importance of scalar reasoning in discourse.

6. The logic of ambiguous superlatives

The logic of the inverted polarity items sheds light on an old paradox in the logic of scalar reasoning. Fauconnier (1975b) noticed that with certain verbs of perception and discovery, it sometimes appears that the same propositional frame can support inferencing from either end of a scale. Thus, it seems that both *the loudest noise* in (22a) and *the faintest noise* in (22b) can be paraphrased with *any noise* (22c).

- (22) a. Eve didn't hear even the faintest noise.
b. Eve didn't hear even the loudest noise.
c. Eve didn't hear any noise.

This is a strange result. If scalar reasoning in general depends on the association of a given scale with a propositional function, it should not be the case that the same scale (in this case a scale of noises ranked in terms of loudness) associated with the same propositional function (in this case *Eve didn't hear X*) can give rise to inferences in either direction on a scale. Otherwise, it would seem, chaos would ensue.

The reason chaos doesn't ensue, of course, is that although the sentence frame is the same in (22a) and (22b), there are actually two very different kinds of meanings involved: (22a) points to the absence of anything for Eve to hear, while (22b) is effectively a comment on Eve's inability to hear. Both sentences are emphatic, but they emphasize very different propositions. The paradox here reflects a basic fact about mental states in general, and acts of perception in particular. As Croft (1993: 64) notes, "a mental state is actually a two-way causal relation," depending both on an experiencer's ability to focus on a given stimulus and on the stimulus's potential to impinge on the experiencer's consciousness.

The two propositions associated with the sentence frame *Eve didn't hear X* thus reflect different facets of the complex causal relation which is hearing. The inferences in (22b) depend on a scale ranking experiencers in terms of their ability to perceive a noise. The basic logic is that since smaller things are harder to perceive than bigger things, if someone doesn't perceive even the biggest, most perceptible thing, then presumably they won't perceive anything at all. The inferences in (22a), on the other hand, depend on an "existential scale" (cf. Hoeksema & Rullmann 2001) ranking stimuli in terms of their likely existence. The logic here is the logic of a canonical scale: if larger entities (e.g. louder noises) occur in a given locale, then presumably smaller entities (softer noises) will be found there as well. So if one can't hear even the tiniest noise, one concludes that there must not be any noise at all.

As an anonymous reviewer points out, expressions which evoke these two types of scale behave quite differently in a number of respects. First, not surprisingly, superlatives on existential scales work well in existential constructions, but superlatives marking extreme perceptual abilities cannot.

- (23) a. There wasn't even the faintest noise to be heard.
b. *There wasn't even the loudest noise to be heard.

On the other hand, "perceptual ability" scales regularly support quantificational superlatives at both ends of the scale, depending on the polarity of the proposition. Thus while (22b) emphasizes a failure of auditory perception, (24a) emphasizes auditory acuity. Existential scales, however, typically only work in scale reversing contexts. (24b) cannot be understood as the opposite of (22a) as the use of *even the loudest noise* will not convey that 'every possible noise' was to be heard.

- (24) a. Eve heard even the faintest noise.
b. ??Eve heard even the loudest noise.

The reason such an interpretation is not available, presumably, is that while one can imagine a minimal possible noise, there is no theoretical limit on how loud a noise can be, and so a phrase like *the loudest noise* is naturally interpreted as denoting a specific loud noise rather than the generic loudest noise possible.

Finally, as Hoeksema and Rullmann (2001) point out, the two types of scale behave differently with different scalar adverbs. Thus while both existential and perceptual ability scales allow modification by *even*, only the existential scale allows modification with the more finicky *so much as* construction.

- (25) a. Eve didn't hear so much as the faintest noise.
b. *Eve didn't hear so much as the loudest noise.

These tests clearly show that although the examples in (22a-b) may appear to involve a single propositional schema, they in fact express two different propositions with two very different sorts of scalar logic.

The important point, once again, is that the ordering of elements in a scalar model depends crucially on the role those elements play in an expressed proposition. The trick here is that a single linguistic frame turns out to be the host for two very different propositional schemas with two very different sorts of scalar logic. The ambiguity arises here because of the inherently ambiguous, bicausal nature of perception, which depends both on the salience of a perceptible object and on the acuity of a potential perceiver. This explains why such scalar ambiguities are so rare, and it predicts that they should arise only with a very limited class of predicates. As Fauconnier (1975b) notes, this is indeed the case: one tends to find such dual readings systematically and exclusively with verbs of perception and discovery—verbs like *see*, *detect* and *find* in (26-28).

- (26) a. I can't see the most distant connection between your argument and your conclusion.
b. With this telescope you can see the most distant galaxies.
(27) a. She couldn't detect the slightest emotion on his face.
b. This device can detect even the slightest facial movement
(28) a. Holmes didn't find the tiniest shred of evidence to support his theory.
b. Holmes can find even the tiniest clues left at a crime scene.

These predicates all refer to situations which depend crucially both on the existence of something to be perceived, and on the ability of some thing or person to perceive them. The scalar logic of these examples just depends on which one of these two aspects is at issue in any given sentence.

7. Conclusions

This paper began with the question how polarity items could exist in a functionally motivated grammar. Answering this question has led us well beyond the confines of grammar narrowly construed, and into the realms of reasoning and rhetoric. In a certain sense, polarity sensitivity is a purely linguistic phenomenon: there is no obvious analogue in any other cognitive domain to the peculiar distributions of polarity items. But the grammar of polarity sensitivity is grounded in a general cognitive ability—the ability to reason in terms of scalar orderings—which manifests itself widely in common-sense reasoning and in the rhetoric of everyday discourse.

This, in any case, is what the Scalar Model suggests: polarity sensitivity is fundamentally a sensitivity to scalar reasoning. Scalar reasoning plays a pervasive role in the structure of rhetorical utterances generally, and polarity items reflect the conventional exploitation of scalar reasoning and complex scalar models for specific rhetorical purposes in discourse. Polarity items exist, therefore, because they are useful, and they illustrate the way grammar tends to harness general conceptual abilities to specific communicative purposes.

Given such a tidy result, it may be well to note that there is still much that needs to be explained. Most importantly, there seems to be more to the distribution of polarity items than can be explained by any single theory of syntax, semantics or pragmatics alone. This is evident just in the fact that different polarity items within the same language, and even very similar polarity items across languages may have subtly, or even radically different sensitivities. A particularly dramatic illustration of this can be found in the rich variety of distributions which Haspelmath (1997) reports for indefinite pronouns in his forty language sample. Ultimately, the fine-grained diversity of polarity sensitivities suggests that their distributions may not depend on general principles alone, but might also reflect speakers' detailed knowledge of what counts as a conventional way of speaking (cf. van der Wouden 1997, Hoeksema 2000 for similar suggestions).

Still, there can be no doubt that polarity sensitivity is subject to some very general principles, and that these principles reflect the inferential properties of polarity contexts. Since Ladusaw (1979), it has been widely assumed that this dependency should be, or indeed must be, formulated in terms of logical entailments, and that these entailments should be captured at a single level of logico-semantic representation in sentence grammar. It is worth noting that Fauconnier in fact drew very much the opposite conclusion: based on quite similar observations, he argued that given the general nature of scalar reasoning and its particular relevance to polarity phenomena, the constraints on polarity items should not be formulated in terms of linguistic representations at all, but rather reflect more general processes of meaning construction and interpretation (Fauconnier 1976, 1978). The evidence I have presented in this paper clearly supports Fauconnier's basic vision. The inferences relevant to polarity licensing need not be, and frequently are not, logical at all—that is, they do not depend on semantic entailment alone, and they cannot be captured at any single level of representation; rather, they seem to depend on a general ability for scalar reasoning which is often used to license conclusions which may be logically invalid but pragmatically quite sound. Polarity items, it seems, are governed by the same sort of inferencing which determines the rhetoric of scalar emphasis and the interpretation of quantificational superlatives, and this inferencing is essentially pragmatic.

Michael Israel
English Department
3101 Susquehanna Hall
University of Maryland
College Park, MD 20742

israel@eva.mpg.de

Acknowledgements

This paper was written while I was a staff scientist in the Dept. of Psychology at the Max Planck Institute for Evolutionary Anthropology in Leipzig. I would like to thank Michael Tomasello for the opportunity to pursue my work in such a pleasant and stimulating intellectual environment. I also thank Bill Croft, Jack Hoeksema, Larry Horn, Paul Kay, Pierre Larrivée, George Lakoff, Tess Wood and an anonymous reviewer for comments, discussion and criticism which helped to clarify my thinking in writing this paper. They have helped make this a better paper than it might have been. The faults which remain are entirely my own responsibility.

References

- Anscombre, Jean-Claude and Oswald Ducrot (1983), *L'Argumentation dans la Langue*. Mardaga, Bruxelles.
- von Bergen, Anke & Karl von Bergen (1993), *Negative Polarität im Englischen*. Narr, Tübingen.
- Croft, William (1993), 'Case marking and the semantics of mental verbs'. In J. Pustejovsky, (ed.) *Semantics and the Lexicon*, pp. 55-72. Kluwer, Dordrecht.
- Dowty, David (1991), 'Thematic proto-roles and argument selection'. *Language* 67.3: 547-619.
- Ducrot, Oswald (1973), *La preuve et le dire*. Maison Mame, Paris.
- Fauconnier, Gilles (1975a), 'Polarity and the scale principle'. *CLS* 11. 188-99.
- Fauconnier, Gilles (1975b), 'Pragmatic scales and logical structures'. *Linguistic Inquiry* 6.3: 353-75.
- Fauconnier, Gilles (1976), *Etude de certains aspects logiques et grammaticaux de la quantification et de l'anaphore en Français et en Anglais*. Atelier Reproduction des Thèses, Lille (1980).
- Fauconnier, Gilles (1978), 'Is there a linguistic level of logical representation?' *Theoretical Linguistics* 5.1: 31-49.
- Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor (1988), 'Regularity and idiomaticity in grammatical constructions: the case of *let alone*'. *Language* 64: 501-38.
- Gaetone, David (1971), *Etude Descriptive du Système de la Négation en Français Contemporain*. Librairie Droz, Genève.
- Giannakidou, Anastasia (1997), *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.
- Haspelmath, Martin (1997), *Indefinite Pronouns*. Clarendon Press, Oxford.
- Heim, Irene (1984), 'A Note on Negative Polarity and Downward Entailingness'. *NELS* 14, 98-107.
- Hirschberg, Julia Bell (1985), *A Theory of Scalar Implicature*. PhD dissertation, University of Pennsylvania.
- Hoeksema, Jack (1994), 'On the Grammaticalization of Negative Polarity Items'. *BLS* 20: 273-82.
- Hoeksema, Jack (1997), 'In days, weeks, months, years, ages: a class of temporal negative polarity items.' Unpublished MS. Groningen University.

- Hoeksema, Jack (2000), 'Negative polarity items'. in L. Horn & Y. Kato, (eds.), *Negation and Polarity: syntactic and semantic perspectives*. pp. 115-146. Oxford University Press, Oxford.
- Hoeksema, Jack & Hotze Rullmann (2001), 'Scalarity and Polarity: a study of scalar adverbs as polarity items'. In V. Sanchez-Valencia and T. van der Wouden, eds., *Perspectives on Negation and Polarity Items*, pp. 129-171. John Benjamins, Amsterdam/Philadelphia.
- Horn, Laurence R. (1972), *On the Semantic Properties of Logical Operators in English*. PhD dissertation, distributed by IULC, 1976.
- Horn, Laurence R. (1989), *A Natural History of Negation*. University of Chicago Press, Chicago and London.
- Horn, Laurence R. (2000), 'Pick a Theory (not just any theory): Indiscriminatives and the free choice indefinite.' In L. Horn & Y. Kato, eds., *Negation and Polarity: syntactic and semantic perspectives*, pp. 147-192. Oxford University Press, Oxford.
- Horn, Laurence R. (2001), 'Flaubert triggers, squattive negation, and other quirks of grammar'. In V. Sanchez-Valencia and T. van der Wouden, eds., *Perspectives on Negation and Polarity Item*. pp. 171 - 98. John Benjamins, Amsterdam/Philadelphia.
- Israel, Michael (1995), 'Negative Polarity and Phantom Reference'. *BLS* 21, 162-173.
- Israel, Michael (1996), 'Polarity Sensitivity as Lexical Semantics'. *Linguistics & Philosophy*. 19: 619-666.
- Israel, Michael (1997), 'The Scalar Model of Polarity Sensitivity'. in D. Forget et al., (eds.) *Negation and Polarity: Syntax and Semantics*. pp. 209-229. John Benjamins, Amsterdam/Philadelphia.
- Israel, Michael (1998), *The Rhetoric of Grammar: Scalar Reasoning and Polarity Sensitivity*. PhD dissertation, Linguistics Department, U.C. San Diego.
- Israel, Michael (1999), 'Some and the Pragmatics of Indefinite Construal'. *BLS* 25.
- Israel, Michael To appear. *Polarity and Scalar Reasoning: The Pragmatic Basis of Grammar*. Cambridge University Press, Cambridge.
- Jackson, Eric (1994), *Negative Polarity, Definites under Quantification and General Statements*. PhD Dissertation, Linguistics Department, Stanford University.
- Kadmon, Nirit & Fred Landman (1993), 'Any.' *Linguistics and Philosophy* 16: 353-422.
- Kay, Paul (1990), 'Even.' *Linguistics and Philosophy* 13: 59-111.
- Kay, Paul (1997), *Words and the Grammar of Context*. CSLI, Stanford.
- Krifka, Manfred (1992), 'Some Remarks on Polarity Items.' In D. Zaefferer, (ed.) *Semantic Universals and Universal Semantics*. 150-189. Foris, Dordrecht.
- Krifka, Manfred (1995), 'The Semantics and Pragmatics of Polarity Items'. *Linguistic Analysis*, 25: 209-257.
- Ladusaw, William (1980), *Polarity Sensitivity as Inherent Scope Relations*. Garland Publishing, New York & London.
- Ladusaw, William (1983), 'Logical Form and Conditions on Grammaticality'. *Linguistics and Philosophy* 6: 373-92.
- Lakoff, George, and Mark Johnson (1980), *Metaphors we live by*. University of Chicago Press, Chicago and London.
- Lakoff, Robin (1969), 'Some Reasons Why There Can't be Any some-any Rule'. *Language* 45: 608-15.

- Langacker, Ronald W. (1987), *Foundations of Cognitive Grammar vol. I: Theoretical Prerequisites*. Stanford University Press, Stanford.
- Langacker, Ronald W. (1991), *Foundations of Cognitive Grammar, volume II: Descriptive Application*. Mouton de Gruyter, Berlin and New York.
- Larrivé, Pierre (1996), 'A Semantic Definition of Negative Polarity Items with Evidence from French and English.' Unpublished MS., Université of Laval.
- Lee, Young-Suk & Laurence Horn (1994), 'Any as Indefinite plus *Even*.' Unpublished MS., Yale University.
- Linebarger, Marcia (1987), 'Negative Polarity and Grammatical Representation'. *Linguistics and Philosophy* 10: 325-87.
- Linebarger, Marcia (1991), 'Negative Polarity as Linguistic Evidence'. *Papers from the Parasession on Negation. CLS* 27. 165-188.
- Michaelis, Laura A. (1993), "'Continuity' within Three Scalar Models: The Polysemy of Adverbial *still*". *Journal of Semantics* 10: 193-237.
- Progovac, Ljiljana (1994), *Negative and Positive Polarity*. Cambridge University Press, Cambridge.
- Rullmann, Hotze (1996), "Two Types of Negative Polarity Items." *NELS* 26: 335-50.
- Schwenter, Scott A. (1999), 'Two types of scalar particles: Evidence from Spanish'. In J. Gutiérrez-Rexach & F. Martínez-Gil (eds.) *Advances in Hispanic linguistics*, pp. 546-561. Cascadia Press, Somerville, MA.
- Schwenter, Scott A. and Shravan Vasishth (2000), 'Absolute and Relative Scalar Particles in Spanish and Hindi'. *BLS* 26.
- Talmy, Leonard (1985), 'Force Dynamics in Language and Thought'. *CLS* 21: 293-337.
- van der Wal, Sjoukje (1996), *Negative Polarity Items & Negation: Tandem Acquisition*. Groningen Dissertations in Linguistics 17, Groningen.
- van der Wouden, Ton (1997), *Negative Contexts: Collocation, polarity and multiple negation*. Routledge, London and New York.
- von Fintel, Kai (1999), 'NPI-Licensing, Strawson-entailment, and context-dependency'. *Journal of Semantics* 16: 1-44.
- Zwarts, Frans (1996), 'A Hierarchy of Negative Expressions'. In H. Wansing, (ed.) *Negation: A Notion in Focus*, pp. 169-194. Walter de Gruyter, Berlin and New York.