In *Negation: Syntax and Semantics*. pp. 209-230. D. Forget, P. Hirschbühler, F. Martineau & M-L. Rivero (eds.). Amsterdam/Philadelphia: John Benjamins. 1997.

The Scalar Model of Polarity Sensitivity: the case of the aspectual operators

Michael Israel
U.C. San Diego

1. Introduction

In recent years theoretical studies of polarity sensitivity have grown increasingly sensitive to the fact that polarity items, both within and across languages, come in a rather bewildering variety of forms. Work by Haspelmath (1993), von Bergen & von Bergen (1993), van der Wouden (1994), Giannakidou (1994), and Rullmann (1996), among others, has given new urgency to the old observation (cf. Horn 1970) that different polarity items often show very different sorts of sensitivities, and one may now wonder not just how we might explain the existence of these sensitivities, but whether indeed there can be any single explanation for such a diversity of forms. Van der Wouden and Rullmann, in particular, are notably pessimistic in this regard, arguing that while polarity items as a class may share basic distributional sensitivities, the reasons for these sensitivities may be quite various. In this paper I take a contrary view, suggesting that while polarity items come in a variety of forms, polarity sensitivity itself reflects a common conceptual schema which unites these forms as scalar operators.

That many polarity items are scalar in nature is hardly subject to doubt. The majority of both negative and positive polarity items (NPIs and PPIs) seem to be transparently quantificational, referring directly to notions of amount, degree, or intensity. In previous work (Israel, in press), I have built on the insights of Fauconnier (1975a,b, 1980) and Kay (1990) to argue that polarity items as a class are scalar operators and that their peculiar distributions reflect their particular scalar semantics. Along with other semantically based accounts (e.g. Ladusaw 1980; Hoeksema 1986; van der Wouden

1994), this work views polarity sensitivity in terms of constraints on the inferential properties of the contexts which license polarity items. Also, in line with recent work by Krifka (1991, 1994), Kadmon & Landman (1993), and Lee & Horn (1995), the work seeks to explain these constraints directly in terms of lexical semantic properties of the polarity items themselves. The goal, ultimately, is to find a conceptual basis which might explain not only why polarity items behave in the ways they do, but also why polarity items, as a class, should exist in the first place.

I pursue this goal here by examining a set of polarity items--the English aspectual adverbs *already, yet, still* and *anymore*--which are not so transparently quantificational and so which might seem to challenge the view that polarity sensitivity is essentially a matter of scalar semantics. The basic puzzle for these forms is how to specify their semantic contributions in a way that will both capture the logical relations which unite them as a class, and explain the constraints on their distributions with respect to things like aspect, aktionsart and polarity. While I can hardly offer a complete solution to this puzzle, I do hope to solve one important piece thereof. I will argue that the reason these forms are sensitive to polarity is that, like other polarity items, they are scalar operators and must be construed with respect to a properly constructed scalar model.

To do this, of course, I must first show why polarity should be considered a scalar phenomenon in the first place, and what it means for a polarity item to be a scalar operator. It is to these basic questions that I now direct my attention.

2. Polarity and the Scalar Principle

The first problem with polarity items is to explain why they are licensed in certain contexts and not in others. This is the licensing problem. As is well known, polarity items are sensitive to a wide range of triggers besides simple negation. These include, among others, questions, conditionals, the standard of a comparative, the restriction of a universal quantifier, complements of adversative predicates like *be surprised* and *doubt,* and various weakly negative operators such as *few, seldom* and *rarely*. The question is, what could possibly make such a motley array of triggers a coherent natural class? In

this paper I adopt the solution suggested by Fauconnier (1975 a,b, 1980): that polarity triggers are united in terms of the sorts of scalar inferencing they allow.

Fauconnier's basic insight came with the observation that polarity items are not alone in their sensitivity to polarity. Superlative NPs, it turns out, can sometimes receive quantificational readings as implicit universals, but these are sensitive to the same sorts of contexts that affect polarity items. In (1a), for example, under negation, *the simplest puzzle* gets a quantificational reading not available in the positive context of (1b).

(1) a. Norm can't solve the simplest puzzle. (=He can't solve any puzzle)
b. +Norm can solve the simplest puzzle. (≠ He can solve any puzzle)

The examples in (2) show that it's not just negation that triggers this effect: other standard licensers of NPIs give similar results.

(2) a. Can Norm solve the simplest puzzle?
b. I'm surprised that Norm can solve the simplest puzzle.
c. Few of the students could solve the simplest puzzle.
d. Everyone who could solve the simplest puzzle got a prize.

In all these examples, the superlative can be roughly paraphrased by the determiner *any*. In (2a), for example, the question may not be just whether Norm can solve the one puzzle which happens to be the easiest, but indeed whether he can solve any puzzle at all. Similarly in (2d), if folks are rewarded merely for solving the simplest puzzle, presumably they also get prizes for solving any of the harder ones.

The sensitivity of superlatives to these contexts provides a clue as to what might unite them. Intuitively, a superlative designates an extreme value in a scalar ordering of elements: thus, *the simplest puzzle* marks an endpoint on a scale in which puzzles are ranked in terms of their difficulty. Scales in general are associated with a sort of common sense logic. If someone can solve a difficult puzzle, we assume that they can also solve any easier puzzle. Conversely, if someone cannot solve a simple puzzle, we assume that they can't solve any harder ones either. And if someone can't solve even the

simplest puzzle, then presumably they can't solve any of them at all. Apparently then, the quantificational readings in (1-2) require contexts in which assertions about a simple puzzle can license inferences about all of the harder puzzles.

Fauconnier, building on Horn's (1972) work on semantic scales, developed the notion of a pragmatic scale to model this sort of obvious but non-trivial inferencing. The basic idea has since been further refined in the work of Fillmore, Kay and O'Connor (FKO 1988) and Kay (1990) on scalar models.

A pragmatic scale is an ordered set of elements associated with a propositional schema. Loosely, a scalar model is just the set of propositions derived from such an association, where the propositions are ordered in a way that supports inferences between them. The model consists of a propositional schema *P*(x, y, ...), with one or more free variables, and for each variable a set of possible values ordered along some semantic dimension. Figure 1 gives a simple, one dimensional example. The model consists of an open proposition, *P*, "Norm can solve y," and a set of possible values for y ordered along a dimension of difficulty.

A Scalar Model of Puzzles

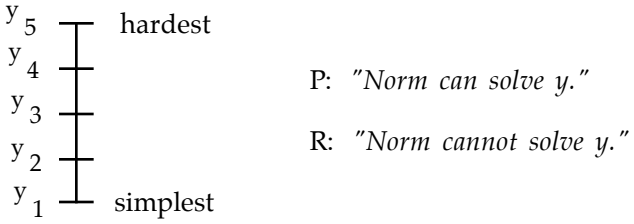

Figure 1

In general, if Norm can solve a given puzzle, then he should be able to solve all the easier ones: the model with schema *P* thus allows for inferences from high propositions to propositions ranked lower in the scale. These inferences are not necessarily valid outside of a scalar model, but they are practically, if not logically, valid and so Fauconnier refers to them as "pragmatic entailments". Fauconnier (1975a,b) captures the logic of scales

in his scale principle: given a schema *P* and values y_1 and y_2 on a pragmatic scale, if $y_2 > y_1$, then $P(y_2)$ will pragmatically entail $P(y_1)$.

Of course, not all scalar models license inferences in the same direction. As we've seen, in a negative sentence--with a schema like *R*, "Norm cannot solve y,"--pragmatic entailments are reversed, running from low to high scalar values : the inability to solve a simple puzzle implies an inability to solve any harder ones. And, of course, negation is not the only context which reverses entailments in a scalar model. Questions, conditionals, the standard of a comparative and the restriction of a universal quantifier, to name a few, all produce similar effects. As Fauconnier (1980) shows in detail, the set of contexts which share this property of reversing scalar implications is in fact precisely the set of contexts which license NPIs. So perhaps this is the solution to the licensing problem: the acceptability of a polarity item in any given context depends on the scalar inferential properties of that context.

But why should polarity items be sensitive to the direction of entailments in a scalar model? In some cases the answer is simple. Just like the quantificational superlatives, many NPIs literally designate a scalar endpoint. Some of these, like *the foggiest notion* and *in the slightest,* are themselves superlatives indicating minimal degrees; others, like *sleep a wink, lift a finger,* and *a shred of evidence,* simply designate a stereotypical minimal unit on some scale. As NPIs, these forms are like superlatives which only allow a quantificational reading: they have no inherent referential value, but they are acceptable where they can be used emphatically, as a way of triggering reference to an ordered set of elements (Israel 1995:164). And of course this is only possible where pragmatic inferences are licensed from low scalar values to higher ones.

So, at least for the minimizer NPIs, the sensitivity to scalar inferencing seems intuitively well-motivated. But how should this sensitivity be represented in the lexicon? And more importantly, how can we extend this sort of intuitive explanation to other polarity items with similar sensitivities but with apparently rather different scalar semantic properties? I suggest that polarity items as a class are sensitive to patterns of scalar inferencing because they are all scalar operators. The ability to construe a situation relative to a scalar model is, I claim, a fundamental aspect of human conceptual structure. Scalar operators are simply forms which impose such a construal on their

expressed content: that is, they presuppose a scalar model available in the context, and they require the information they express to be integrated with that scalar model in a particular way (cf. FKO 1988, Kay 1990). More precisely, I argue that polarity items conventionally encode two sorts of properties inherent in the construction of any scalar model: quantitative value and informative value. It is to these that I now turn.

3. Two Scalar Properties and Four Sorts of Polarity Item

Every proposition within a scalar model can be distinguished on the basis of two basic properties: the first of these, q (or *quantitative*) value, refers to a proposition's position within a scalar ordering: the higher a proposition is along a scale, the higher its q-value will be; the second, i (or *informative*) value, refers to a proposition's relative informativeness within the model: the more propositions a given proposition pragmatically entails, the higher is its i-value.

Since any scalar model is based on some ordering, the notion of q-value should be unproblematic. Following Rivara (1990), Koenig (1993), and others, I distinguish the dimensions along which gradable properties are ordered from the scales that actually order them. In Koenig's terms, a dimension is an ordered set of sets: the dimension of size thus consists of sets of individuals such that all the elements in a given set have the same size and each of the ordered sets corresponds to a unique size. Scales are more complicated: they take a dimension and pair it with one of two converse orderings, or 'polarities'. Thus for the size dimension we have two scales, which we can arbitrarily call positive and negative: the positive, or big scale is ordered in terms of increasing size, from little things to big things; the negative, or little scale has the converse ordering, from big things to little things. The words *big* and *little* themselves refer to some vague range of elements towards the high end of their respective scales. And more generally, the q-value of any proposition in a scalar model reflects its position as high or low along the scale that defines the scalar model.

Of course, what counts as high or low on a scale depends on background assumptions and implicit norms: what's big for a mouse may be quite small for a house. I assume that in context the use of a scalar predicate always evokes some scalar norm as

an implicit standard of comparison. The particular value of the norm depends on the expectations and assumptions of the speech act participants, but in general, it simply reflects a default, real world understanding of whatever is under discussion. The existence of the norm allows us to view the essentially gradient notion of q-value as a simple binary opposition: a proposition has a high q-value if it lies above the scalar norm and a low q-value if it lies below the scalar norm within a scalar ordering.

The usefulness, and indeed necessity of the scalar norm is yet more striking when we consider i-value. I-value depends on a proposition's inferential relation to other propositions in a model, but how is this relation determined and with respect to which other propositions? If, as seems reasonable, the scalar norm constitutes an essential, if unspoken, aspect of any scalar model, then i-value can be understood directly in terms of an asserted proposition's inferential relation to the norm. The norm, in effect, constitutes a presupposition as to what proposition within a model would, in some default context, be most likely to be asserted. Basically then, what it means for an assertion to be construed with respect to a scalar model is that it is implicitly contrasted with some alternative default proposition. Following Kay (1990), I will refer to the asserted proposition overtly encoded by a sentence the *text proposition*, or TP, and I will call the contrasting proposition defined by the scalar norm the *context proposition*, or CP. In general, if an asserted TP entails the CP, then the TP has a high i-value, and if the TP is (or would be) itself entailed by the CP, then the TP has a low i- value.

Again, by defining i-value relative to a single implicit CP we reduce a gradient notion to a binary distinction: propositions entailing the scalar norm have a high i-value; propositions entailed by the norm have a low i-value. This should be intuitive. In general, if a proposition entails the norm, its assertion is informative because it exceeds what one might normally expect. Such relatively informative propositions I call *emphatic*. On the other hand, if a proposition is itself entailed by the norm, then its assertion will be distinctly uninformative, for it will leave uncertain whether or not the default expectation of the norm is met as well. I call these under-informative propositions *understating*.

Thus far we have characterized i-value and q-value as properties of propositions. Essentially, what it means for a lexical item to encode one of these properties is that the

proposition to which it contributes must be construed relative to the the structured set of alternatives in a scalar model: i-value and q-value thus do not simply add information to a proposition, rather, they situate a proposition within a sort of informational matrix. In general, if a form conventionally encodes either an i-value or a q-value, it counts as a scalar operator and must be interpreted relative to a scalar model. But if a form encodes both an i-value and a q-value, it will also be a polarity item: the combination of a fixed scalar location (q-value) with a fixed inferential relation to the scalar norm (i-value) limits a form to just those contexts in which the direction of scalar inferencing is simultaneously compatible with both of its scalar values. Thus, the minimizer NPIs discussed above combine a low (in fact, a minimal) q-value with a high (or emphatic) i-value, and this combination constrains their distribution to contexts where low scalar values are more informative than high scalar values.

Since q-value and i-value are both, effectively, binary features, we should find four basic types of polarity items based on their combinations. The minimizers offer a clear example of one of these basic types, combining low q-values with high i-values. NPIs like the English *all that* (as in "he's not all that late") and *much* ("he didn't waste much time") illustrate a second group, in which high q-values combine with low i-values. As NPIs, these forms show roughly the same distributional constraints as the minimizers, though their pragmatic purpose in life is quite the opposite, being used not to strengthen but rather to mitigate or hedge the force of a negative utterance.

Positive polarity items show a similar division into emphatic and understating forms. Emphatic PPIs include quantificational idioms like *heaps of, scads of* and *the whole shebang,* and degree modifiers like *awfully* and *utterly*. These forms encode high scalar q-values in highly expressive, high i-value assertions: they effectively signal a high degree of speaker commitment to an expressed proposition. Understating PPIs also include quantificational idioms and degree modifiers--*a tad, a smidgen*, *pretty, fairly,* and *kinda,* among others. These forms encode (relatively) low scalar q-values in hedged, low i-value assertions: their use tends to signal either a certain tentativeness, or perhaps a desire not too insist too strongly on one's point. (For further discussion of the four polarity types, both in English an other languages, see Israel (in press)).

In forms which are conventionally specified for both, q-value and i-value conspire to create polarity sensitivity. Consider the distribution of the NPI *sleep a wink* in (3).

(3) a. Marianne didn't sleep a wink that night.
b. *Marianne slept a wink that night.

Here, the emphatic NPI contributes a minimal scalar value to an expressed TP and further requires that the TP be more informative than a CP based on the scalar norm. In (3a), since the TP, 'M didn't sleep the smallest amount', entails the CP, 'M didn't sleep a normal amount', the requirement is met: the NPI counts as emphatic and the sentence is well-formed. In (3b), however, the NPI cannot properly express its emphatic i-value: here the expressed TP, 'M slept the smallest amount', is itself entailed by the CP, 'M slept a normal amount.' This produces an understatement, and so renders the emphatic NPI unacceptable.

Similar considerations apply to the understating PPI *a smidgen* in (4).

(4) a. Brandon had a smidgen of jelly on his collar.
b. *Brandon didn't have a smidgen of jelly on his collar.

As a scalar operator, *a smidgen* contrasts with a set of alternative scalar values in a scalar model. The expressed proposition in (4a), 'that B had a slight amount of jelly on his collar,' is construed relative to some 'normal' CP implicit in the evoked model, something like 'B had a moderate amount of jelly on his collar.' (Note that the CP does not require any particular expectation that B should have jelly on him; rather, given that he is so besmirched, the scalar norm just reflects what might be expected to be a normal smirch in such circumstances.) Since having a moderate amount of jelly on one's collar entails having a slight amount there, the CP is more informative than the TP. The PPI happily expresses its low i-value, and so the sentence is understating and grammatical. In (4b), however, the implication reversal triggered by negation makes the understating *a smidgen* unacceptably emphatic. Here the TP, that 'B didn't have a slight amount...,'

entails the implicit CP, that 'B didn't have a moderate amount ...' The emphatic effect is at odds with the understating i-value of the PPI, and so the sentence is at best peculiar.

As these examples suggest, the particular combinations of q-value and i-value in PPIs are such that they are only compatible with contexts where inferences run from high to low q-values; contrariwise, the particular combinations in NPIs are such that they are compatible only with implication reversing contexts, where inferences run from low to high q-values. This, in essence, is why polarity items are sensitive to polarity. Polarity in general is a matter of scalar inferencing and polarity items are just scalar operators: the proper expression of their lexical semantics depends on the availability of a properly constructed scalar model. This explanation seems compelling at least for the sorts of examples discussed so far. Still, if polarity sensitivity in general really is just a matter of scalar semantics, similar stories will have to be told for those polarity items which are not so transparently quantificational in nature. In the next section I consider one important set of these: the aspectual operators *still, already, yet* and *anymore*.

4. Aspectual Polarity Items

Aspectual adverbs like the English *already, yet, still* and *anymore* may make modest contributions to sentence meaning, but they have inspired a voluminous literature (Horn 1970; König 1977, 1991; Löbner 1987, 1989; Mittwoch 1988; Garrido 1992; Michaelis 1992, 1993; van der Auwera 1993; *inter alia*). While this work has greatly illuminated the lexical semantics of these forms, the explanation of their polarity sensitivity remains somewhat murky. At best, the problem tends to be solved by means of an ad hoc suppletion rule (cf. Traugott & Waterhouse 1969), or else simply ignored. For my part, I can offer scant insight into the many complex issues raised by these forms and the literature they have inspired; I can, however, offer a simple explanation for their polarity sensitivity in terms of their lexical semantics.

In what follows, I follow the many who have followed Horn (1970) in analyzing these forms as marking relations between two phases of a given eventuality, one presupposed and the other asserted. Basically, my proposal is just that these phases represent distinct propositions within a scalar model. Michaelis (1993) has already

analyzed *still* as a scalar operator. I suggest that all four of these forms are scalar operators, and furthermore, that their polarity sensitivity is a reflection of their scalar semantics. I propose that the four are related as operators on two sorts of scales defined by the two converse orderings of the time dimension: *already* and *yet* operate on inceptive scales ordering states by how early they begin; *still* and *anymore* operate on continuative scales, in which states are ordered by how late they end.

All four operators involve the evaluation of an asserted proposition (the TP) with respect to a presupposed scalar norm or background expectation. Each situates the TP with respect to some scalar norm, and each requires a specific inferential relation between the TP and an implicit CP associated with the norm. In other words, these forms, like their more transparently quantificational counterparts, are conventionally specified for both q-value and i-value: as such, as with other polarity items, they are limited to just those contexts in which both features can be felicitously expressed.

4.1. *already* and *yet:* Inceptive Scales. Despite years of scrutiny, there is still little agreement on the semantics of *already*. As Michaelis puts it, the only clear consensus is that *already* is "a sentential adverb which encodes or at least reflects the existence of a given state of affairs [the 'already' state or AS] at reference time [RT]" (1992: 322). Beyond this, any analysis must somehow account for two basic intuitions about the 'already' state. First, that the AS is understood as holding somehow earlier than expected: it only makes sense to say that something is already the case if it wasn't expected until later. Second, that the AS seems to contrast with some negative phase before RT: it only makes sense to say that something is already the case if one thought it might not have been the case before. I propose to capture these intuitions by analyzing *already* as an operator on what I will call inceptive scales.

An inceptive scale is a scale in which propositions denoting states are ordered in terms of the earliness of their onsets: the earlier a proposition holds, the higher its q-value. As such, a proposition of the form [already P] is roughly glossed by 'P as early as RT'. In simple affirmative contexts inceptive scales license inferences from early to later propositions: early propositions with high q-values entail less early ones with lower q-values, as shown by the examples in (5a). As with any scale, implication reversing

contexts, like negation in (5b), reverse the direction of these entailments so that low q-value later propositions entail high q-value earlier propositions.

(5) a. Huey was drunk as early as noon. — —>
Huey was drunk as early as sundown.
b. Huey wasn't drunk as early as noon <— —
Huey wasn't drunk as early as sundown.

As a scalar operator, *already* encodes a high q-value on an inceptive scale. This explains why the AS seems so early: a*lready*'s q-value marks the asserted TP as early with respect to some scalar norm. The norm here represents a default expectation about the likely potential onset for the 'already' state coded by the TP. As in any scalar model, the norm defines an alternative, context proposition against which the TP is evaluated. The fact that the AS is understood as holding earlier than expected thus simply reflects the fact that *already* ranks its TP high on a scale of propositions ranked by earliness.

Already's high q-value also helps explain why the AS is normally understood as contrasting with a prior negative phase: if the AS were known to hold in the first place, it would seem less than relevant, if not downright incoherent, to assert its holding at RT as being relatively early. I follow Michaelis (1992) in viewing the prior negative phase as a non-conventional implicature associated with the word's use: the essence of *already* is just priority to expectation on an inceptive scale, but the assertion of such priority is normally only relevant where the anticipated state has just come into existence.

This implicature is also facilitated by one final aspect of *already:* the TP must not only precede the context proposition on a scale of earliness, it also has to be understood as exceeding it. In other words, there is a sense in which the use of *already* signals that the asserted 'already' state goes beyond what was expected. Garrido (1992: 367) thus suggests that *already* and *still* (or rather, their Spanish equivalents, *ya* and *todavia)* each make an assertion against the background of a contrary expectation; in a similar vein, Van der Auwera speaks of a counterfactual course of events against which *already* is evaluated (1993: 621). Within the scalar framework adopted here, the same intuitions are captured by having *already* encode a high i-value, requiring the asserted TP to be more

informative than an implicit CP lower on the scale. The high i-value combined with the already posited high q-value also explains *already*'s polarity sensitivity: together, the two features insure that *already* can only appear in contexts where higher scalar values pragmatically entail lower ones--in other words, positive polarity contexts.

According to Michaelis *already* "codes the existence of a state of affairs at a reference time, prior to a reference interval containing a like type" (1992: 321). This is wholly compatible with the analysis here: that *already* situates an asserted TP as holding high on an inceptive scale at RT, prior to an implicit, and less informative, CP.

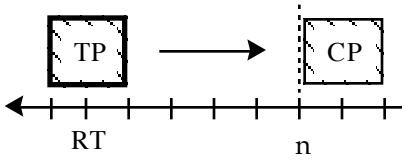

Figure 2: *already*

Figure 2 provides a schematic illustration. Since by convention time flows from left to right, the arrow on the left of the time line indicates that earlier times are ranked higher than later ones. The dashed line, n, before the CP indicates the normal or expected onset of a state, and the arrow from the TP to the CP indicates that the former pragmatically entails the latter. The figure presents *already* as combining a high q-value with a high i-value on an inceptive scale: as with other polarity items, the combination of the two limits *already* to just those contexts where both can be happily expressed.

To see how this works, consider the examples in (6) (adapted from Michaelis 1992).

(6) a. Huey was already drunk at noon.
b. *Huey wasn't already drunk at noon.

In (6a), the use of *already* suggests that Huey's noontime inebriation was earlier than one might have expected, though maybe not in itself surprising. What (6a) asserts, the TP, is that 'Huey was drunk as early as noon.' *Already* contrasts this assertion with a set of alternative propositions situating the onset of inebriation at points lower on the

scale. Note that the scalar norm does not depend on any particular expectation that Huey would be drunk; rather, the expectation is simply that if one is to be drunk, there are more normal times later in the afternoon. The norm supplies an implicit CP--something like 'Huey was drunk as early as [5 p.m.]'--representing a less remarkable time for Huey (or anyone) to be drunk. Since if Huey was drunk by noon, he was *a fortiori* drunk by 5, the TP here entails the CP. *Already*'s high i-value is thus expressed in an emphatic assertion and the sentence is grammatical. In (6b), however, the implication reversal triggered by negation prevents the emphatic effect. Here the TP, that 'Huey was not drunk as early as noon,' is itself entailed by the CP, that 'Huey was not drunk as early as [5 pm].' The expressed proposition is thus understating, in conflict with *already*'s emphatic i-value, and so the sentence is starred.

Since *yet* and *already* are at least apparent suppletives, I assume they are denotationally equivalent: both encode high q-values on inceptive scales and both can be glossed 'as early as [RT]'. Like *already*, *yet* marks a situation at RT as holding prior to a situation of the same sort on an inceptive scale. The difference between them is a matter of construal, reflecting the fact that they encode contrary i-values: *already* has a high i-value requiring an asserted TP to be construed with respect to a weaker alternative CP; *yet* has a low i-value requiring its TP to contrast with a stronger alternative CP. As such, *yet* contributes to a proposition which is compatible with, rather than contrary to, some default expectation. In figure 3, the CP marks the last phase of a state before an expected change and is understood as entailing the TP asserted to hold at RT.

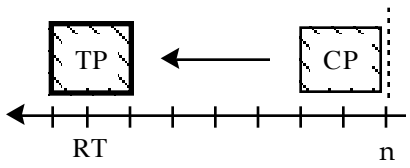

Figure 3: *yet*

On this account *yet* is an understating NPI analogous to forms like *much* and *all that*: it situates an asserted TP high on an earliness scale, and it presupposes the

availability of some more informative proposition lower down the scale (i.e. later) which might just as well have been asserted. Consider the examples in (7).

(7) a. At midnight Angela hadn't stopped talking yet.
b. *At midnight Angela had stopped talking yet.

In (7a), the use of *yet* suggests that Angela's loquacity did not last forever: an end must have come, perhaps no later than dawn. This expectation of an end supplies a CP representing the latest (i.e. the least early) phase at which one might expect her talking to continue. Since negation reverses entailments on an inceptive scale, the later CP, that 'A hadn't stopped ...as early as dawn,' entails the asserted TP, that 'A hadn't stopped...as early as midnight'. This conforms to the schema in figure 3: the asserted TP is less informative than implicit CP, and so the sentence is understating and well-formed.

In (7b), again, the use of *yet* suggests that the asserted YS, here that 'A had stopped', would not last indefinitely. Again, the expectation supplies an implicit CP marking the last phase at which the YS might be anticipated to hold. But here the CP, that 'A had stopped as early as dawn', does not entail the TP, that 'A had stopped as early as midnight.' Quite the reverse. This conflicts with the requirement that *yet* form an understating assertion, and the sentence is consequently quite ungrammatical.

The analysis here posits a subtle distinction between *yet* and *already* in terms of the relative informativeness of their background CPs. Support for this analysis comes from a subtle difference between the two in questions. Because questions are non-monotonic and do not force any scalar inferences in any particular direction, they tend to allow either the NPI *yet* or the PPI *already*. But as noted by Horn (1970:321) and van der Auwera (1993:632), the choice between the two is significant: (8a), with *yet,* poses a neutral question, but (8b), with *already,* is strongly biased towards a positive response.

(8) a. Has Larry read your paper yet?
b. Has Larry read your paper already?

The difference follows from the two forms' distinct scalar semantics. Both sentences presuppose an expectation that Larry will read the paper at some point, and both ask whether this point has been reached. With *already* the question posed is effectively whether the event has been realized earlier than was expected. In other words, the focus of the question is *already* itself, and more particularly its high i-value: the question asks whether the situation at reference time (i.e. the TP) is such that it would entail what is expected to hold later on (the CP). With *yet*, on the other hand, the question posed is effectively whether the event has been realized, not in excess of, but in accordance with what was expected: in other words, is the situation at reference time (i.e. the TP) such that it would be entailed by what is expected later on (the CP)? *Yet* presupposes that the event will take place, but without suggesting that this will happen in any way ahead of schedule: it thus produces a neutral question. *Already* also presupposes that the event will take place, but it further suggests that this might happen earlier than was expected: this results in a distinctly biased question.

Interestingly, the same contrast between a neutral *yet* and a biased *already* shows up in conditionals, which, like questions, tend to allow both NPIs and PPIs. Although *yet* does not often occur in conditionals (cf. Horn 1970:318), under the right conditions it clearly can . The attested example in (9a) was addressed to a new player in a poker game who was anxious to watch a few hands before starting to play herself.

(9) a. "If you're playing yet, it's your deal, otherwise it goes to Josh."
b. If you're playing already, it's your deal.

The use of *yet* here signals an expectation that the addressee will in fact be joining the game, but it remains neutral as to just when this might take place. Contrast this with the less welcoming (9b), where the use of *already* would suggest an expectation that the new player will wait out at least a few more hands . Again, with *yet* the early TP is understood as being compatible with the expectations of the later CP, while with *already*, the emphatic i-value marks the earliness of the expressed TP as exceeding the expectation implicit in the CP.

4.2. *still* and *anymore*: Continuative Scales. Just as *already* and *yet* are operators on inceptive scales, *still* and *anymore* are operators on continuative scales. While inceptive scales rank states according to how early they begin, continuative scales rank states according to how late they last. The two scale types thus involve converse orderings on the dimension of time: inceptive scales give early propositions high q-values; continuative scales give later propositions high q-values. As such, propositions of the form [still P] and [anymore P] gloss roughly as 'P as late as RT'. As with other scales, in simple affirmative contexts a continuative scale licenses inferences from high to low q-values, later propositions entailing earlier ones, as in (10a). As with other scales, implication reversing contexts, as in (10b), reverse the implications, allowing inferences from low, early propositions to higher, late propositions.

(10) a. Alice was chopping garlic as late as noon. — —>
Alice was chopping garlic as late as mid-morning.
b. Alice wasn't chopping garlic as late as noon. <— —
Alice wasn't chopping garlic as late as mid-morning.

The analysis of *still* and *anymore* basically mirrors that proposed for *yet* and *already*. *Still* and *anymore* mark an asserted TP as holding high on a continuative scale and as contrasting with an implicit CP lower on the scale. Again, the difference between the two is simply a matter of contrasting i-values: *still* marks an asserted TP as emphatic and hence as entailing its associated CP; *anymore* marks a TP as understating and hence as being itself entailed by the CP. Figures 4 and 5 provide an informal illustration.

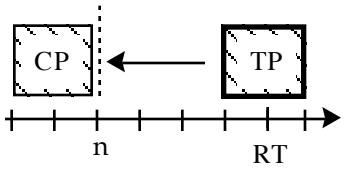

Figure 4: *still*

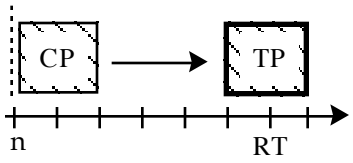

Figure 5: *anymore*

The arrow on the left of the time line indicates that later times are ranked higher in the ordering. The dashed lines associated with the norm (*n*) represent default expectations about the end of a state. In the case of *still*, the expectation is that the state coded by the TP might be over at *n*: the TP holding later and entailing the earlier CP means that this expectation is exceeded. In the case of *anymore,* the expectation is that the converse of the state coded by the TP might be over at *n*: here, the later TP being entailed by the earlier CP means that the TP is compatible with the default expectation.

Again, polarity sensitivity is a function of scalar semantics: in order for its TP to count as emphatic, *still* requires a context allowing inferences from high to low q-values; *anymore,* in order for its TP to count as understating, requires a context in which inferences run from low to high q-values. As the licensing of these forms simply mirrors that of the inceptive operators, I confine myself here to an illustration for *still*.

(11) a. Alice was still cooking at midnight.
TP: Alice was cooking as late as midnight [t = RT]. ——>
CP: Alice was cooking as late as 9 [t < RT].
b. *Alice wasn't still cooking at midnight.
TP: Alice wasn't cooking as late as midnight. —/—>
CP: Alice wasn't cooking as late as 9.

In (11a) the use of *still* suggests that Alice's cooking might have ended at some more reasonable hour, say by 9 p.m. This expectation yields a CP marking the anticipated final phase of cooking, and the later TP thus indicates that the expectation has been exceeded. The TP entails the CP and so, consonant with the scalar semantics of *still,* the assertion counts as emphatic. In (11b), however, negation reverses the relationship between TP and CP: just because Alice wasn't in the kitchen at midnight is no reason to assume she wasn't busy there earlier. Here the assertion does not even equal, let alone exceed the default expectation encoded by the CP. This contradicts *still*'s emphatic semantics and yields an ungrammatical sentence.

Since the analysis here for *still* and *anymore* parallels that of *yet* and *already*, the prediction is that the two sets of operators should also parallel each other in behavior.

Horn (1970) suggests that this is not quite the case and that the contrast in questions between a biased *already* and a neutral *yet* does not show up with *still* and *anymore*.

(13) a. Does Gladys still smoke pot?
b. Does Gladys smoke pot anymore?

The questions in (13) are close to being mutual paraphrases, but, *pace* Horn, they are not entirely interchangeable. (13a) with *still*, for example, seems more appropriate where the speaker would be shocked if Gladys persisted in such passé pastimes (this reading is particularly facilitated by focus accent on *still* itself). *Still*, in other words, suggests, or can suggest, that the text proposition somehow emphatically exceeds normal expectations. The effect of *anymore* is milder. (13b) presupposes that Gladys did smoke pot and may no longer, but it doesn't quite suggest that she shouldn't: the text proposition is thus understood as being compatible with prior expectations.

In general, though the distinction is subtle, *anymore* suggests that a situation will not last indefinitely and may in fact be over, while *still* suggests that a situation has already lasted longer than one would have expected and probably should be over. The examples in (14) help bring this contrast into sharper relief.

(14) a. Are Jack and Jill still married?
b. ?Are Jack and Jill married anymore?

Here s*till* biases the question, suggesting that the marriage was not so durable to begin with and probably should be over by now: the question works well in a society where divorce is not uncommon. The question in (14b), however, is unnatural precisely because it is more neutral: *anymore* does not so much anticipate the demise of Jack and Jill, as it presupposes that marriages in general do not last indefinitely. But despite decades of social upheaval our default understanding is still that marriage does last indefinitely, and the neutral question clashes with this default understanding. Although I cannot explain why the distinction between *still* and *anymore* should be more subtle than

that between *yet* and *already,* I conclude that it is in kind, if not in degree the same: emphatic PPIs yield biased questions, understating NPIs allow for neutral questions.

4.3. Strict Sensitivities and Other Operators. Thus far I have argued that aspectual polarity items are scalar operators just like other polarity items. But the aspectual operators are special: *yet* and *anymore* are strict NPIs, with distributions more restricted than those of other NPIs. Any full account of what makes these forms polarity sensitive should, therefore, have something to say about why they are so extra sensitive. In the space remaining, I can at least make some suggestions.

The examples in (15-16) show one way that the analysis of *yet* and *anymore* as operators on a temporal dimension helps explain their heightened sensitivities. Here we have what seems to be a minimal semantic contrast, but one with clear grammatical consequences: in the (a) sentences *yet* and *anymore* are licensed in the scope of *only,* while the apparent semantic paraphrases in the (b) sentences are ungrammatical.

(15) a. Angela is the only one who's tasted the cookies yet.
b. *Angela is the first one who's tasted the cookies yet.
c. Angela is the first one who ever did a damn thing for me.

(16) a. Narsai is the only one who eats radicchio anymore.
b. *Narsai is the last one who eats radicchio anymore.

The example in (15c) shows that superlatives like *the first one* do normally license NPIs. The problem is that while the exclusive particle *only* does not require a scalar ordering on its contrast set (cf. König 1991), *the first* and *the last* do . In (15b), for example, *yet* situates an asserted TP, something like 'x has tasted the cookies as early as now,' at a particular point in time. At the same time, *the first* requires that alternative values for x be ordered in terms of temporal precedence. In other words, *the first* in this context requires that a set of simultaneous propositions at reference time to be temporally ranked. But this is incoherent, and so the sentence doesn't make sense.

Another way of viewing the problem here might be in terms of the expected change of state which defines the scalar norm for these operators. On this view, the problem in (15b) is that if Angela is the first to taste the cookies, she will remain so forever: no change is possible and so there can be no expectation of a change of state. I assume that something like this, at least, explains a more pervasive restriction preventing *yet* and *anymore* from being licensed by factive adversatives, as in (17-18).

(17) a. *I'm surprised that Narsai has eaten yet.
b. *It's weird that Narsai has eaten yet.

(18) a. *?I'm surprised that Gladys smokes anymore.
b. *?It's weird that Gladys smokes anymore.

In these examples the factivity of the matrix verb seems to clash with the expectations imposed by *yet* and *anymore*. *Yet* presupposes an expectation that a new state may begin in the future, while the factive predicates in (17) require this state to have already begun; *anymore* presupposes that an old state may have ended, while the factives of (18) require that it persists. Though I don't doubt that this is the right intuition, I'm afraid I may not yet have made it precise enough to count as an explanation. Basically, the idea is that as aspectual operators *yet* and *anymore* presuppose not just a properly structured scalar model, but also an expected transition between two phases of a given eventuality. It is this extra, aspectual aspect of their meanings that makes them extra sensitive. While I think this is reasonable, a more fully explicit formulation of the intuition will have to await further study.

Finally, it is worth noting that the analysis here predicts the existence of other sorts of aspectual operators: in addition to emphatic PPIs like *still* and *already,* and understating NPIs like *yet* and *anymore* we should expect to find understating PPIs and emphatic NPIs operating on the same inceptive and continuative scales. Plausible candidates are not hard to come by. *Finally,* for example, appears to be an understating counterpart to the emphatic *already*. As van der Auwera notes (1993:618), while *already* marks a state as beginning relatively early, *finally* marks it as relatively late. In other

words, *finally* situates an expressed proposition low on an inceptive scale and marks it as falling short of some prior expectation: it is thus a low scalar PPI. A similar analysis may apply to the expression *so far* (meaning roughly 'as of yet') as an understating PPI and a counterpart to the emphatic *still* on continuative scales. And, as I have argued elsewhere (Israel 1996), punctual *until* may be usefully understood as a low scalar emphatic NPI on an inceptive scale: it forms a maximally informative proposition by designating the lowest point on a scale of earliness at which a state holds. While I toss these out as quick suggestions here, the viability of such analyses will be an important test for the usefulness of the scalar story I have tried to tell. The suggestion is that constructs like inceptive and continuative scales are basic conceptual building blocks which organize lexical semantic structure: obviously, a demonstration of their principled recurrence in a variety of forms will help make this suggestion much more compelling.

5. Conclusions

There is much more to the aspectual operators than I have been able even to touch on in the space of this paper. The analysis here is intended only to capture those essential elements of their meaning which make them polarity sensitive, and as such, the analysis is perhaps a bit deceptively neat. Each of these forms represents a complex category in its own right; each has its own particular sensitivities and distributional idiosyncracies; and each participates in a range of usages related to, but distinct from the basic senses discussed here. In short, much of what is most interesting about these forms, the nitty-gritty messy details of their usage and lexical semantics, has simply been ignored; nonetheless, I do hope to have captured the essence of their meanings.

While I began this paper by noting the tremendous diversity among polarity items, my goal throughout has been to reveal a basic conceptual unity beneath this diversity. At this point it is worth considering precisely what this means. To claim that *yet* and *anymore* share certain scalar semantic properties with other negative polarity items is not to deny the tremendous differences in form, function and behavior that separate them. What they share is a very abstract sort of meaning, not so much a common semantic content as a common way of construing content. Scalar operators require an expressed

proposition to be interpreted with respect to an ordered set of alternative propositions, but beyond the basic, schematic notions of quantitative and informative value, scalar operators can vary radically in terms of the dimensions on which they build and the kinds of information which they encode.

The claim that all polarity items are scalar operators is thus in some sense a weak claim about polarity items, for many different things can be scalar operators and there is always more to any polarity item than the simple fact that it is a scalar operator. The notions of i-value and q-value and of scalar operators in general simply provide a basic schematic framework for the unified analysis of an extremely multifarious phenomenon. They are the common denominator for a rich variety of lexical forms, but they do not in themselves constrain or explain the richness of that variety. On the other hand, the claim that all polarity items are scalar operators is a strong claim about the nature of scalar inferencing. The claim is that a basic conceptual ability, the ability to reason in terms of a scalar model, plays a central role in the organization of lexical and grammatical information. Polarity sensitivity, on this account, is just the specifically linguistic manifestation of a much broader cognitive phenomenon: it is still a matter of grammar, but a grammar which is itself built on general cognitive abilities.

Acknowledgements

I would like to thank Gilles Fauconnier, Paul Hirschbühler, Ron Langacker, Louise McNally, and Laura Michaelis for comments and encouragement offered while I was working on this paper. Naturally, the defects that remain are entirely my own fault.

References

Fauconnier, Gilles. 1975a. "Polarity and the Scale Principle." *CLS* 11. 188-99.

——. 1975b. "Pragmatic Scales and Logical Structures." *Linguistic Inquiry* 6: 353-75.

——. 1980. *Etude de Certains Aspects Logiques et Grammaticaux de la Quantification et de L'Anaphore en Français et en Anglais*. Dissertation, Université de Paris VII, 1976. Lille: Atelier Reproduction des Thèses.

Fillmore, Charles J., Paul Kay, & Mary Catherine O'Connor. 1988. "Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*." *Language* 64: 501-38.

Garrido, Joaquín. 1992. "Expectations in Spanish and German Adverbs of Change." *Folia Linguistica* XXVI: 357-402.

Giannakidou, Anastasia. 1994. "The Semantic Licensing of Negative Polarity Items and the Modern Greek Subjunctive." In *Language and Cognition 4: yearbook of the group for theoretical and applied linguistics of the University of Groningen.*

Haspelmath, Martin. 1993. *A Typological Study of Indefinite Pronouns*. PhD thesis, Freie Universität Berlin.

Hoeksema, Jack. 1986. "Monotonicity Phenomena in Natural Language." *Linguistic Analysis* 16:25-30.

Horn, Laurence R. 1970. "Ain't it Hard Anymore." CLS 6: 318-327.

——. 1972. *On the Semantic Properties of Logical Operators in English*. PhD dissertation, distributed by IULC, 1976.

Israel, Michael. 1995. "Negative Polarity and Phantom Reference." *BLS 21*.

——. 1996. "Until 2." paper presented at the conference, *Conceptual Structure, Discourse and Language 2*. SUNY, Buffalo.

——. In press. "Polarity Sensitivity as Lexical Semantics." to appear in *Linguistics & Philosophy*.

Kadmon, Nirit & Fred Landman. 1993. "Any." *Linguistics and Philosophy* 16:353-422.

Kay, Paul. 1990. "Even." *Linguistics and Philosophy* 13: 59-111.

Koenig, Jean-Pierre. 1993. "The lexical semantics of scalar predicates: a unified account." Paper presented at the annual meeting of the LSA. Los Angeles.

König, Ekkehard. 1977. "Temporal and Non-Temporal Uses of 'noch' and 'schon' in German". *Linguistics and Philosophy* 1, 173-198.

——. 1991. *The Meaning of Focus Particles*. London: Routledge.

Krifka, Manfred. 1990. "Some Remarks on Polarity Items." In D. Zaefferer (ed.), *Semantic Universals and Universal Semantics*. 150-189. Dordrecht: Foris.

——. 1994. "The Semantics and Pragmatics of Weak and Strong Polarity Items in Assertions." *Proceedings of SALT IV,* 195-219.

Ladusaw, William. 1980. *Polarity Sensitivity as Inherent Scope Relations*. New York & London: Garland Publishing.

Löbner, Sebastian. 1987. "Quantification as a Major Module of Natural Language Semantics." In J. Groenendijk et al. (eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. 53-85. Dordrecht: Foris.

——. 1989. "German *Schon - Erst - Noch:* An Integrated Analysis." *Linguistics and Philosophy* 12:167-212.

Lee, Young-Suk & Laurence Horn. 1995. "*Any* as Indefinite plus *Even."* Unpublished Ms. Yale University.

Michaelis, Laura. 1992. "Aspect and the Semantics-pragmatics interface: the case of *already*." *Lingua* 87: 321-339.

——. 1993. "'Continuity' within Three Scalar Models: The Polysemy of Adverbial *Still*." *Journal of Semantics* 10:193-237.

Mittwoch, Anita. 1988. "Aspects of English Aspect: on the interaction of perfect, progressive, and durational phrases." *Linguistics and Philosophy* 11: 203-54.

Rivara, René. 1990. *Le Systéme de la Comparaison*. Paris: Editions de Minuit.

Rullmann, Hotze. 1996. "Two Types of Negative Polarity Items." *NELS* 26.

Traugott, Elizabeth & John Waterhouse. 1969. "*Already* and *Yet:* a suppletive set of aspect markers?" *Journal of Semantics* 5: 287-304.

van der Auwera, Johan. 1993. "*Already* and *Still*: Beyond Duality." *Linguistics and Philosophy* 16: 613-653.

van der Wouden, Ton. 1994. *Negative Contexts*. Groningen Dissertations in Linguistics: University of Groningen.

von Bergen, Anke & Karl von Bergen. 1993. *Negative Polarität im Englischen*. Tübingen: Narr.