The Semantics and Pragmatics of Polarity Items

Manfred Krifka

For some thirty years negative polarity items (NPIs) have provided crucial evidence for linguistic theory. But the various accounts of NPIs have not yet attained explanatory adequacy. The goal of this paper is to derive the distribution of polarity items (and in particular of different types of polarity items) from their semantic structure and independently motivated pragmatic principles.

Section (1) provides an overview of existing theories of NPIs and their problems. In section (2), I outline my explanation of the distribution of so-called weak polarity items, and in (3) I discuss the semantic nature and distribution of strong polarity items. Section (4) offers a comparison of weak and strong NPIs. Section (5) discusses a wider range of polarity items. Section (6) is devoted to so-called „double licensing“, and section (7) to certain locality effects. In section (8) I will discuss NPIs in questions.

1. Polarity Items: Past Theories, Current Problems

1.1. Syntax and Semantics

There is an ongoing debate between syntacticians and semanticists about the proper explanation of the distribution of NPIs. Klima (1964) may be seen as the earliest proponent of a syntactic theory. According to him, NPIs must be „in construction with“, or in more recent terms, be c-commanded by, a trigger. Triggers are either an overt negation or an „affective element“, e.g. a verb like *surprised*.

(1) a. John didn't say anything.

b. We were surprised that John said anything.

Baker (1970) reduced the set of triggers to negation, eliminating „affective elements“. He claimed that NPIs may be licensed derivatively by semantic entailment. For example, a sentence like (1.b) entails a sentence in which the NPI anything is licensed directly, by negation.

(2) (1.b) entails: We expected that John wouldn't say anything.

A problem of this theory, already observed by Baker, is that it may very well predict that NPIs occur everywhere, as every sentence ϕ will entail $\neg\neg\phi$. Indirect licensing therefore must involve a more specific semantic relation than logical entailment.

Fauconnier (1975, 1978, 1980) and Ladusaw (1979) approached the issue directly from a semantic angle by claiming that NPIs occur in downward-entailing (DE) contexts and denote extreme elements among a set of alternatives. A downward-entailing context for α, i.e. an expression XαY, is defined as a context where replacing α with a semantically weaker constituent β yields a stronger expression XβY. For example, the nominal argument of a universal determiner is DE (cf. 3.a,b), and consequently it allows for NPIs (cf. 3.c). Here I use „$\subseteq$" to express the relation of semantic strength; $\alpha \subseteq \beta$ means that α is at least as strong (or specific) as β.

(3) a. carrots $\subseteq$ vegetables

b. Everyone who ate vegetables got sick. $\subseteq$ Everyone who ate carrots got sick.

c. Everyone who ate any vegetables got sick.

This line of attack was undermined by Linebarger (1980, 1987, 1991), who observed that many NPI contexts are not really DE. For example, the protasis of conditionals allows for NPIs like ever (cf. 4.c) but fails to show the DE property (cf. 4.a,b), contrary to claims made by Ladusaw.

(4) a. You visit China and get sick there. $\subseteq$ You visit China.

b. If you visit China you will enjoy it. $\not\subseteq$
If you visit China and get sick there you will enjoy it.

c. If you ever visit China you will enjoy it.

Linebarger also showed that adversative predicates, like *surprised*, are not DE in Ladusaw's sense. Furthermore she pointed out that NPIs have to occur in the immediate scope of their licenser, which seems to call for a syntactic analysis. She illustrated this constraint with quantifiers and reason clauses. For example, in the following sentence only the narrow-scope reading for *every party* is possible (i.e., there are no earrings that Mary wore to every party), not the wide-scope reading (i.e., it was not to every party that Mary wore earrings).

(5) Mary didn't wear any earrings to every party.
a. Not (Some earrings x (Every party y (Mary wear x to y)))
b. *Not (Every party y (Some earrings x (Mary wear x to y)))

Heim (1987) defended the semantic position by showing that the notion of DEness may be restricted in an intuitively appealing way. Essentially, she claims that the presence of NPIs signals DEness along a scale specified by the NPI and with respect to a particular position in a sentence. For example, she analyzed *ever* as meaning 'at least once', and having alternatives meaning 'at least n times', where n>1. The protasis of a conditional sentence like (4.c) exhibits this limited DEness, as the following implication holds:

(6) a. *ever*: 'at least once'; alternative values: 'at least n times', n>1.

b. If you visit China at least once you will enjoy it. $\subseteq$
If you visit China at least n times, you will enjoy it. (n>1).

Another important innovation is that Heim makes the acceptability conditions of NPIs dependent on the current common ground of the conversation. For example, (6.b) is not meant to be a logical truth, but a truth of suitable common grounds at which sentence (4.c) can be uttered felicitously; (6.b) can be seen as a presupposition of (4.c). Hence NPIs are not just passive elements that may or may not be licensed: they actively accomodate common grounds.

More recently, at least three interesting approaches have been developed that deserve more careful examination: Progovac (1988, 1992, 1993) tries to explain the distribution of NPIs by binding theory, Kadmon & Landman (1993) propose an account of *any* in terms of semantic strength, and Zwarts (1993) develops an algebraic theory that distinguishes between different NPI types and contexts.

1.2. The Binding-Theoretic Approach of Progovac

In her dissertation (1988) and several articles (1992, 1993), Progovac points out the following problem:

(7) a. Mary didn't remember anything.

b. Mary forgot that anyone came yesterday.

c. *Mary forgot anything.

The standard semantic account of NPIs can deal with (7.a) and (7.b): In (a), the NPI is licensed by overt negation, and in (b), by the negation inherent in *forget*, which can be paraphrased as 'not know anymore'. However, the ungrammaticality of (7.c) then constitutes a problem, as one of its readings can be paraphrased by 'not know anymore' as well, cf. e.g. *Mary forgot this poem*.

Progovac proposes that NPIs must be licensed either by negation or by an operator „Op“ in the specifier position of the same clause. Thus, NPI licensing turns out to be subject to principle A of binding theory, the principle that governs, among other things, the distribution of reflexives (cf. Chomsky 1981). The operator Op in turn is semantically restricted: It can occur only in clauses that are not upward-entailing. This proposal can be illustrated with the following examples:

(8) a. Mary didn't remember anything.

b. Mary forgot [Op that anyone came yesterday]

c. *(OP) Mary forgot anything.

d. *Mary remembered [(Op) that anyone came yesterday].

In (8.a) the NPI *anyone* is licensed by clausemate negation. In (b) it is licensed by an operator Op that can occur here, due to the fact that the clausal argument of *forget* is not upward-entailing: If Mary forgot that a woman came yesterday, she might not have forgotten that a person came yesterday. (c) is out: There is no overt negation, and the operator Op cannot occur either, as root sentences are upward entailing. For example, if Mary forgot a poem by Goethe, then she forgot a poem, but not necessarily vice versa. Also, a case like (d) is out, as the clausal argument of *remember* is upward-entailing: If Mary remembers that a young woman came yesterday, she remembers that a woman came yesterday, but not necessarily vice versa.

The crucial piece of evidence for Progovac is the contrast between (8.a) and (c). Progavac seems to make the prediction that NPIs can never be licensed in the non-clausal argument position of a non-negated root clause. However, this is not the case, as the following examples show:

(9) a. John lacks any sense of humor.

b. John came without any present.

I think examples like (9.a,b) show that the binding-theoretic analysis of NPIs is on the wrong track. Within a semantic analysis the difference between verbs like *forget* and verbs like *lack* can be accounted for by assuming that *forget* (in the non-clausal version) has an object position of type e, whereas *lack* and prepositions like *without* have objects of the quantifier type $\langle\langle e,t\rangle,t\rangle$.[1] This is corroborated by the fact that the latter ones, but not the former, allow for non-specific readings of indefinite NPs:

(10) a. John lacks a place to live. [some place or other]

b. John came without a coat. [some coat or other]

c. John forgot a poem by Goethe [a specific poem].[2]

Now assume that NPIs like *anything* are of type $\langle\langle e,t\rangle,t\rangle$, just like other nonreferring NPs. Consequently, *anything* can stay in situ as an object of *lack* or *without*, but must take scope over the predicate as an object of *forget*. The meanings of the predicates in question contain a negation; we may paraphrase *lack* as 'not have', and *forget* as 'not know anymore'. Then we see immediately that *anything* is licensed in (11.a,b), but not in (c), as in this case *anything* is outside of the scope of forget.

(11) a. lack anything: $\lambda x.\mathbf{lack}(x,\mathbf{anything})$

b. come without anything: $\lambda x.\mathbf{without}(\mathbf{anything})(\mathbf{come})(x)$

c. forget anything: $\lambda x.\mathbf{anything}(\lambda y.\mathbf{forget}(x,y))$

Another argument Progovac adduces for the binding theoretic account for NPIs is that there are languages which seem to have NPIs that can only occur in the immediate scope of a clausemate negation. Progovac cites English *until*[3] and negative terms in Serbo-Croatian:

(12) a. John did not arrive until 7 o'clock.

b. *I do not claim that John arrived until 7 o'clock.

c. *It is not the case that John arrived until 7 o'clock.

(13) a. Milan ne voli nikoga
Milan not loves noone

b. *Ne tvrdim da Milan voli nikoga.
not I-claim that Milan loves noone

The situation of Serbo-Croatian is quite widespread; it obtains in all languages that exhibit negative concord (cf. Ladusaw 1992). I agree with Progovac that such examples show that the expressions in question have local cooccurrence restrictions with a negative element. But I would like to reserve the term „Negative Polarity Item“ for expressions like *anything* whose distributions are not directly dependent on the occurrence of a clausemate negation. Negative concord can be described as a grammaticalized agreement, a distinct phenomenon. Note that phrases that show negative concord all contain a negative element, which is *ni-* in Serbo-Croation and other Slavic languages, whereas typical negative polarity items do not contain negation elements.[4]

1.3. The Strengthening Approach of Kadmon & Landman

Kadmon & Landman (1993) deal only with NPIs based on the determiner *any*. According to their theory, such NPIs are used to indicate a reduced tolerance to exceptions, or, in other words, a widening of the extension of an indefinite NP. This is taken to be a lexical property of *any*: It is said that *any* is licensed only if the widening that it induces creates a stronger statement (their principle C). They illustrate this with the following example; assume that speaker A asks speaker B (a cook for a group of 50 people):

(14) A: Will there be French fries tonight?
B: No, I don't have potatoes.
A: Maybe you have just a couple of potatoes that I could fry in my room?
B: Sorry, I don't have ANY potatoes.

According to Kadmon & Landman's description, B had the impression that his first answer was misunderstood in a way that *potatoes* is interpreted as 'enough potatoes for the whole group'. In his second answer, the use of *ANY potatoes* indicates that potatoes has to be understood in a wider sense than before.

Kadmon & Landman offer interesting and convincing solutions for a range of apparent counterexamples to Ladusaw's theory. For example, they point out that adversative predicates like *be surprised* are indeed downward-entailing once a certain perspective is fixed. They describe the occurrence of *any* in the protasis of conditionals as a widening of implicit restrictions. And they propose a theory of free-choice *any* as involving a marking of NPs in the restrictor of a generic statement.

But there are also problems with their analysis. First, it seems that *any* expresses widening only when it is stressed. Notice that B's first answer in (14) could have been *No, I don't have any potatoes,* where it is implausible that any widening is intended, and that B's second answer requires stress on *ANY*. Kadmon & Landman argue that it is not stress, but the presence of *any* that induces widening, but their reasons are not wholly convincing (cf. also Rohrbaugh 1993).

A second problem is that NPIs based on *any* can be used in contexts where the notion of reduced tolerance to exceptions is problematic. For example, we can say, referring to a particular sequence of numbers: *This sequence doesn't contain any prime numbers*. It seems implausible that *any prime numbers* induces a widening of the precise concept 'prime number' here, or even a contextual widening from 'small prime number' to 'small or large prime number'.

Third, a semantic rule like Kadmon & Landman's (C) is problematic for theoretical reasons as it refers in the semantic description of one expression to the larger context in which this expression is used, and hence is intrinsically non-compositional. We may grant

(C) the status of a descriptive generalization, but the next question should be: At which level is (C) checked, and what is responsible for this checking?

1.4. The Algebraic Theory of Zwarts

Zwarts (1993) takes serious an earlier observation by various authors (e.g., Horn 1978, Edmondson 1981) that not all NPIs are equal. Zwarts identifies three classes of NPIs which he calls „weak“, „strong“, and „superstrong“, and gives an algebraic characterization of the contexts that can host these different types of NPIs.

Weak NPIs, like *need*, *care* and presumably unstressed *any* and *ever* just require that the context in which they occur is monotone decreasing, or DE. Phrased in functional terms, a context f is monotone decreasing iff it holds that $X \subseteq Y$ entails $f(Y) \subseteq f(X)$. We find such NPIs, for example, in the scope of quantifiers like *few students* or *fewer than three students*.

(15) a. Few students have ever gone to the library.

b. Fewer than three students cared to hand in a paper.

c. At most five students have gained any financial support.

Strong NPIs, like *any student at all*, or *lift a finger*, *bat an eyelash* etc. need a context that, in addition to being DE, has the property of being „anti-additive“. A context f is anti-additive iff $f(X \cup Y) = f(X) \cap f(Y)$, where $\cup$ and $\cap$ are Boolean disjunction and conjunction. A quantifier like *fewer than three students* is not anti-additive, in contrast to a quantifier like *no student*:

(16) a. Fewer than three students smoked cigarettes or drank beer. ≠
Fewer than three students smoked cigarettes and fewer than three students drank beer.

b. No student smoked cigarettes or drank beer. =
No student smoked cigarettes and no student drank beer.

Consequently, we find contrasts like the following one:

(17) a. *Fewer than three students {lifted a finger/read any book at all}.

b. No student {lifted a finger / read any book at all}.

The reported judgements follow Zwarts (1993). I have found that English speakers in general see a grammaticality difference between sentences like (17.a) and (b), but they are unlikely to judge sentences like (17.a) as strictly ungrammatical.

Superstrong NPIs, for which Zwarts gives the Dutch example *mals* (lit. 'tender, soft') and the English example *one bit*[5], can only occur in a context that is downward-entailing, anti-additive and satisfies the condition $f(\neg X) = \neg f(X)$, where „$\neg$“ expresses generalized negation or complementation; Zwarts calls these contexts „anti-morphic“. A quantifier like *no student* does not satisfy this condition, but sentential negation does:

(18) a. No student wasn't happy. ≠ It is not the case that no student was happy.

b. John wasn't happy. = It is not the case that John was happy.

Consequently, we find contrasts such as:

(19) a. John wasn't one bit happy about these facts.

b. *No linguist was one bit happy about these facts.

Although Zwarts' study is a very important contribution that adds considerable refinement to our understanding of NPIs, it has some empirical problems and leads to new theoretical challenges:

First, the distinction between the three classes of polarity items is less clear than suggested by Zwarts. Various NPIs classified as weak by Zwarts, like *hurt a fly*, seem to be rather of the strong type.

Second, there seems to be an interesting relation between NPI types and stress that Zwarts does not mention and that does not follow straightforwardly from his analysis: As a general rule, weak NPIs are unstressed, whereas strong NPIs attract stress.[6] This can be seen in the contrast between weak *any* and strong *any (whatsoever)* and its Dutch and German equivalents *ook maar iets* and *auch nur irgendetwas*:

(20) a. No child got {any presents /ANY presents (whatsoEVer)}.

b. Fewer than three children got {any presents /*ANY presents (whatsoEVer)}.

Third, the conditions of monotone decrease and anti-additivity are not sufficient for Zwarts' purposes, as they would be satisfied by a function f that maps every set X to a specific element. Examples are quantifiers like *zero or more students* or *some arbitrary number of students*, which always yield a true sentence when combined with a VP. However, these quantifiers do not license NPIs, neither strong ones nor weak ones.

Another problem is that the class of superstrong NPIs doesn't seem to be definable in terms of anti-morphicness, or in any algebraic terms for that matter. If it were, we should not find any contrast between the following examples, contrary to the facts:

(21) a. John wasn't one bit happy about these facts.

b. *It is not the case that John was one bit happy about these facts.

It seems that Zwarts' class of superstrong NPIs coincides with Progovac's class of NPIs that have to be licensed by a clause-mate negation (where negation need not be restricted to standard negation, but may include emphatic negation, such as German *keineswegs* or Dutch *allerminst*). Therefore I will disregard this class in the present article, for the same reason as I disregarded negative concord phenomena.

A more general point is: Why do different types of NPIs require different types of contexts? Why does the distribution of weak and strong NPIs seem to depend on algebraic concepts like monotone decrease or anti-additivity? In this paper I will address this very question: Why is it that certain types of polarity items only occur in certain contexts? I will propose that this is due to a peculiar interaction between the meaning of polarity items and the expressions in which they occur, and certain general pragmatic rules that come with the illocutionary force of the sentence. The theory of polarity items proposed here is an elaboration of ideas presented first in Krifka (1990, 1992).

2. The Semantics and Pragmatics of Weak NPIs

In this section I will develop the theory I am going to propose with a simple example: licensing of the NPI *anything* in an assertion in the scope of negation. As indicated above, the explanation will have two parts, involving the semantics of polarity items and the pragmatics of the sentences in which they occur.

2.1. The Semantics of Weak NPIs

The basic assumptions concerning the semantics of NPIs like anything are: (a) NPIs introduce alternatives; and (b) the alternatives induce an ordering relation of semantic specificity, where the NPI itself denotes a most specific element in that order.

According to (a), NPIs resemble items in focus as viewed by focus theories such as Rooth (1985, 1992). I will incorporate alternatives using structured meanings which have been developed to capture the semantic impact of focus (cf. Jacobs 1984, von Stechow 1990). More specifically, I will use triples $\langle B,F,A \rangle$, where B stands for the background, F for the foreground (the polarity item or the item in focus), and A for the set of alternatives to F. The set of alternatives A contains items of the same type of F, but not F itself[7]. Typically, when B is applied to F, we will get a standard meaning B(F).

Semantic strength, rendered by „$\subseteq$“, is defined for all types based on the truth-value typet as follows:

(22) a. If α, β are of type t, then $\alpha \subseteq \beta$ iffÿ20$\alpha \to \beta$.

b. If α, β are of type $\langle \sigma,\tau \rangle$, then $\alpha \subseteq \beta$ iff for all γ of type σ: $\alpha(\gamma) \subseteq \beta(\gamma)$.

For example, if P, Q are properties (type $\langle s,\langle e,t \rangle\rangle$), then $P \subseteq Q$ iff $\forall i \forall x[P(i)(x) \to Q(i)(x)]$. Thus, we have **sparrow**$\subseteq$**bird**, as the set of sparrows is a subset of the set of birds in all possible worlds i. As usual, we will write $\alpha \subset \beta$ iff $\alpha \subseteq \beta$ and $\neg\beta \subseteq \alpha$, and say that α is „stronger“ than β.

Let me introduce an example. The NPI anything is analyzed as the following BFA-structure:

(23) *anything*: $\langle$B, **thing**, {P| P$\subset$**thing**}$\rangle$

Here, **thing** is the most general property (a notion that depends on the context and on selectional restrictions in ways that are not accounted for here). The precise nature of the background B is a function of the syntactic position in which *anything* occurs, e.g. as object or subject. The alternatives are a set of properties that are stronger than the most general property, **thing**. For simplicity of exposition I will assume that every property that is more specific than **thing** is an alternative. In any case, one important requirement for the set of alternatives is that it is exhaustive in the sense that all the alternatives together make up the foreground.

(24) Exhaustivity requirement: $\cup\{P| P\subset\mathbf{thing}\} = \mathbf{thing}$

I will now derive the meaning of two sentences in which *anything* occurs in object position. In order to do so we have to work with interpretation rules that can handle structured meanings. Assume that we already have rules that give ordinary, non-structured interpretations, then structured meanings can be integrated as follows:

(25) a. If a semantic rule calls for application of α to β,
and $\beta = \langle B,F,A\rangle$,
then $\alpha(\beta) = \alpha(\langle B,F,A\rangle) = \langle\lambda X[\alpha(B(X))],F,A\rangle$,
where X a variable of the type of F.

b. If a semantic rule calls for application of α to β,
and $\alpha = \langle B,F,A\rangle$,
then $\alpha(\beta) = \langle B,F,A\rangle(\beta) = \langle\lambda X[B(X)(\beta)],F,A\rangle$,
where X is a variable of the type of F.

These rules guarantee that information about the position where the focus is interpreted and about the alternatives is projected from daughter nodes to mother nodes.[8]

Now let us derive the meaning of a sentence that will yield a bad assertion, *Mary saw anything*. I assume a semantic representation language with explicit reference to possible worlds; in general, if α is a constant of type $\langle s,\tau\rangle$, then α_i, short for $\alpha(i)$, is the extension of α at world i. I will write R(x,y) for R(y)(x). The semantic combination rules are functional application, modulo the provision for BFA structures:

(26) *anything*

$\langle \lambda Q\lambda R\lambda i\lambda x\exists y[Q_i(y) \wedge R_i(x,y)], \textbf{thing}, \{P|\ P\subset\textbf{thing}\}\rangle$,

$= \langle B,F,A\rangle$

|

| *saw*

| **saw**

|/

saw anything

$\langle B,F,A\rangle(\textbf{saw})$,

$= \langle \lambda Q[B(Q)(\textbf{saw})],F,A\rangle$

$= \langle \lambda Q[\lambda R\lambda i\lambda x\exists y[Q_i(y) \wedge R_i(x,y)](\textbf{saw}), F, A\rangle$

$= \langle \lambda Q\lambda i\lambda x\exists y[Q_i(y) \wedge \textbf{saw}_i(x,y)], \textbf{thing}, \{P|\ P\subset\textbf{thing}\}\rangle$

|

| *Mary*

| $\lambda P\lambda i[P_i(\textbf{m})]$

|/

Mary saw anything

$\langle \lambda Q\lambda i\exists y[Q_i(y) \wedge \textbf{saw}_i(\textbf{m},y)], \textbf{thing}, \{P|\ P\subset\textbf{thing}\}\rangle$

We get a BFA-structure with a B component that, when applied to F, will yield the proposition $\lambda i\exists y[\textbf{thing}_i(y) \wedge \textbf{saw}_i(\textbf{m},y)]$, i.e. the set of worlds i where Mary saw something.

A sentence like *Mary didn't see anything* can be analyzed, somewhat simplified for expository reasons, as involving a negation operator applied to the BFA-structure we arrived at above:

(27) | NEG,

| $\lambda p\lambda i[\neg p_i]$

|/

Mary didn't see anything,

$\langle \lambda Q\lambda i\neg\exists y[Q_i(y) \wedge \textbf{saw}_i(\textbf{m},y)], \textbf{thing}, \{P|\ P\subset\textbf{thing}\}\rangle$

When we apply the B component to F, we get the proposition $\lambda i\neg\exists y[\textbf{thing}_i(y) \wedge \textbf{saw}_i(\textbf{m},y)]$, the set of worlds i in which Mary saw nothing.

I would like to point out an important fact that will be crucial for the following discussion. In both cases (26) and (27) we obtained a BFA structure that defines a proposition, B(F), and a set of alternative propositions, $\{p|\exists F'[F'\in A \wedge p=B(F')]\}$. And as we have a certain logical relationship between the foreground F and its alternatives F' (F being weaker than any alternative F'), we have a certain logical relationship between B(F) and its alternatives B(F'). In the case of (26) B(F) is weaker than any alternative proposition B(F'): The set of worlds where Mary saw something or other is a proper superset of every set of worlds where Mary saw something that is described in more specific terms. In the case of (27) B(F) is stronger than any alternative proposition, as the set of worlds where Mary didn't see anything is a proper subset of the set of worlds where Mary didn't see something that is described in more specific terms. Hence we can say that the logical relationship between F and its alternatives is „preserved“ in the semantic compositions that lead to (26), but it is „reversed“ in the semantic composition with negation that leads to (27). In both cases we may say that the BFA structure is „projected“.

So much for the semantic part of the story. The question now is, why is (26) bad, but (27) good? I propose that the reason for this is to be found in pragmatics, in particular, in the felicity conditions for assertions.

2.2. The Pragmatics of Standard Assertion

Let us adopt the following, rather standard theory of assertions (cf. Stalnaker 1972):

a) The participants of a conversation assume, for every stage of the conversation, a mutually known common ground c. For our purposes we can represent common grounds as sets of possible worlds.

b) If one participant asserts proposition p, and the audience does not object, the current common ground c is restricted to $c \cap p$. We may assume certain felicity conditions, e.g. that $c \cap p \neq c$ (that is, p expresses something that is not already established), and that $c \cap p \neq \emptyset$ (that is, p doesn't express something that is taken to be impossible). I will say that p is „assertable“ with respect to the common ground c in this case.

We may stipulate an assertion operator **Assert** that, when applied to a proposition, takes an input common ground c to an output common ground $c \cap p$:

(28) **Assert**($\langle$B,F,A$\rangle$)(c) = $c \cap$B(F), iff B(F) is assertatble w.r.t. to c and

a) For all $F' \subseteq A$ such that $c \cap B(F') \neq c \cap B(F)$:
the speaker has reasons not to assert B(F'),
that is, to propose $c \cap B(F')$ as the new common ground.

b) There are $F' \in A$ such that B(F') is assertable w.r.t. c, and
$c \cap B(F') \neq c \cap B(F)$.

If (a) or (b) are not met, the assertion is undefined. But in general the conditions will trigger accomodation of the common ground.

Condition (a) states that the speaker has reasons for not asserting alternative propositions B(F'). There are various possible reasons — the speaker may know that B(F') is false or lack sufficient evidence for it. One typical case has been described as scalar implicature (cf. Gazdar 1979, Levinson 1984). Example:

(29) Mary earns $2000.

Implicature: Mary doesn't earn more than $2000.

This implicature arises in the following way. Let us assume that $2000 introduces the set A of all alternative amounts of money, e.g.

$A = \{..., \$1998, \$1999, \$2001, \$2002 ...\}$

Then the assertion of (29) can be analyzed as follows, using the previously defined assertion operator; from here on I will generally suppress condition (28.b) for simplicity.

(30) $\mathbf{Assert}(\langle\lambda X\{i|\mathbf{earn}_i(\mathbf{m},X)\},\$2000,A\rangle)(c) = c\cap\{i|\mathbf{earn}_i(\mathbf{m},\$2000)\}$

iff for all $F'\in A$ with $c\cap\{i|\mathbf{earn}_i(\mathbf{m},F')\} \neq c\cap\{i|\mathbf{earn}_i(\mathbf{m},\$2000)\}$:

Speaker has reasons not to propose $c\cap\{i|\mathbf{earn}_i(\mathbf{m},F')\}$.

In the current example the proposition asserted and the alternative propositions stand in a relation of semantic strength to each other: *Mary earns $2000* entails *Mary earns $n*, for *n*<2000, and is entailed by *Mary earns $m*, for 2000<*m*. In such cases we can distinguish two types of reasons the speaker has if he or she wants to be both truthful and informative:

i) If $[c\cap B(F)] \subset [c\cap B(F')]$, the reason is that $[c\cap B(F')]$ would be less informative.

ii) If $[c\cap B(F')] \subset [c\cap B(F)]$, the reason is that the speaker lacks sufficient evidence for proposing $[c\cap B(F')]$ as the new common ground. If the speaker does not indicate otherwise — e.g. by asserting *Mary earns at least $2000*, or *Mary earns $2000 and perhaps more* — the reason is more specifically that the speaker thinks that $[c\cap B(F')]$ is false, and the hearer is entitled to draw this inference.

Of course, (i) is (one part of) Grice's maxim of Quantity, and (ii) is Grice's maxim of Quality (cf. Grice 1975). Notice that Quantity reasons are related to weaker propositions, whereas Quality reasons are related to stronger propositions.

The configuration we find with scalar implicatures is an important subcase of the general assertion rule. This warrants the introduction of a special operator, **Scal.Assert**. Its triggering condition is that the proposition actually asserted and the the alternative assertions are informationally ordered with respect to each other (31.a). And its semantic impact is that all propositions that are semantically stronger than the proposition made are negated (31.b).

(31) a. $\mathbf{Assert}(\langle B,F,A\rangle)(c) = \mathbf{Scal.Assert}(\langle B,F,A\rangle)(c)$,

if for all $F'\in A$: $[c\cap B(F')]\subseteq[c\cap B(F)]$ or $[c\cap B(F)]\subseteq[c\cap B(F')]$

b. $\mathbf{Scal.Assert}(\langle B,F,A\rangle)(c) =$

$\{i\in c|\ i\in B(F) \wedge \neg\exists F'\in A[[c\cap B(F')]\subset[c\cap B(F)] \wedge i\in B(F')]\}$

In a more refined semantic theory the second conjunct in this set would have the status of a conversational implicature.

Let us apply this view of assertion to our NPI examples. They clearly satisfy the condition for scalar implicatures. For the ungrammatical example (26) we get the following result:

(32) $$\mathbf{Scal.Assert}(\langle \lambda Q \lambda i \exists y[Q_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)],\ \mathbf{thing},\ \{P|\ P \subset \mathbf{thing}\}\rangle)(c)$$
$$= \{i \in c|\ \exists y[\mathbf{thing}_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)] \wedge$$
$$\neg\exists P \subset \mathbf{thing}[\{i \in c| \exists y[P_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)]\} \subset \{i \in c| \exists y[\mathbf{thing}_i(y) \wedge$$
$$\mathbf{saw}_i(\mathbf{m},y)]\} \wedge \exists y[P_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)]\}]]\}$$

Notice that the first conjunct — that Mary saw a **thing** — and the second conjunct — that there is no P, P⊂**thing**, such that Mary saw a P — contradict each other. Whenever Mary saw some x that is a **thing**, x will fall at least under some property P that is defined more narrowly. Technically, every input common ground c will be reduced to the empty set.

For the grammatical example (27) we get the following result:

(33) $$\mathbf{Scal.Assert}(\langle \lambda Q \lambda i \neg\exists y[Q_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)],\ \mathbf{thing},\ \{P|\ P \subset \mathbf{thing}\}\rangle)(c)$$
$$= \{i \in c| \neg\exists y[\mathbf{thing}_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)] \wedge$$
$$\neg\exists P \subset \mathbf{thing}[\{i \in c| \neg\exists y[P_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)]\} \subset \{i \in c| \neg\exists y[\mathbf{thing}_i(y) \wedge$$
$$\mathbf{saw}_i(\mathbf{m},y)]\} \wedge \neg\exists y[P_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)]\}]]\}$$

The first conjunct restricts the common ground c to those worlds i for which Mary didn't see a **thing**. The second conjunct is trivially satisfied here, as it holds for no P, P⊂**thing**, that the proposition that Mary didn't see a P is stronger than the proposition that Mary didn't see a **thing**. The difference between our two examples is that in (26) the proposition B(F) is at least as weak as any alternative proposition, whereas in (27) B(F) is at least as strong as any alternative.

It is important to understand the type of this explanation, as it can easily be misunderstood. A sentence like (26) is not simply bad because it would express a very general meaning.[9] There are sentences that do that without being ungrammatical, namely tautolgies like *War is war*. Rather, (26) is bad because it expresses a sentence in which what is said systematically contradicts what is implicated. The assertion made by (26) says that Mary saw something, but the implicatures deny that Mary saw anything in particular.

The explanation why (26) is bad may become clearer when we contrast it with the following sentence, which is good although it expresses the same proposition as (26):

(34) Mary saw something. $\lambda i \exists y[\mathbf{thing}_i(y) \wedge \mathbf{saw}_i(\mathbf{m},y)]$

In contrast to *anything* in (26), *something* in (36) does not introduce any alternatives and hence does not induce any alternative-related implicatures. This seems at odds with a common analysis that says that *something* is a positive polarity item, and assume that positive polarity items work like NPI's except that their scale is reversed. However, I contend that NPs based on *some* are not polarity items at all. The observation about the scope differences in cases like *Mary didn't see anyone* (¬∃) and *Mary didn't see someone*

($\exists\neg$) that have been adduced for the PPI status of someone rather should be explained as a paradigmatic effect induced by Grice's principle of ambiguity avoidance: In case a speaker wants to express the $\neg\exists$ reading the unambiguous form containing *anyone* is preferred. It might very well be that this paradigmatic effect is so strong that it is virtually grammaticalized.

3. The Semantics and Pragmatics of Strong NPIs

In the preceding section we have derived the basic facts about the distribution of the weak NPI anything. In this section I will address the distribution of strong NPIs, for which I take stressed *anything* or *anything at all* as an example.

3.1. The Semantics of Strong NPIs

There is an important difference between the weak and the strong use of anything:

(35) Mary didn't get anything for her birthday.

(36) Mary didn't get ANYthing (at ALL) for her birthday.

(35) just says that Mary got nothing; (36) stresses the fact that Mary didn't even get some minor present for her birthday. This seems to be a fairly consistent property of stressed *anything* and other expressions based on *any*. Kadmon & Landman (1993), who generally investigate stressed *any*, give a wide variety of examples and argue that they involve widening of the extension of the noun meaning to include borderline cases.

To capture cases like (36) we have to assume a slightly different interpretation of *anything* that highlights the special role of borderline cases, and a special type of assertion that carries the implicature expressed by the word *even* in the paraphrase. I propose the following BFA structure for the meaning of strong anything:

(37) *ANYthing*: $\langle \mathrm{B}, \mathbf{thing}, \{\mathrm{P} | \mathrm{P} \subset \mathbf{thing} \wedge \neg \mathrm{min}(\mathrm{P})\} \rangle$

Here, „min“ is a second-order predicate that identifies properties that are applicable to „minor“ entities of a certain dimension (which is left unexpressed here). For example, in (36) the relevant dimension is the class of birthday presents; a Porsche would rank high in that dimension, whereas piece of chewing gum would rank low and probably be considered minor. The use of a predicate „min“ is preliminary; I will give a more satisfying account in section (5.2).

One important requirement for the BFA-structure in (37) is that the alternatives are non-exhaustive in the following sense:

(38) Non-exhaustivity requirement: $\cup\{\mathrm{P} | \mathrm{P} \subset \mathbf{thing} \wedge \neg \mathrm{min}(\mathrm{P})\} \subset \mathbf{thing}$

This is because **thing** can be applied to minor objects to which no predicate P can be applied. I propose that non-exhaustivity is the distinguishing semantic property for strong NPIs.

3.2. The Pragmatics of Emphatic Assertion

Let us come now to the type of assertion we found in (36). I claim that it is the same type of assertion that we find in examples like the following ones that exhibit emphatic focus:

(39) a. Mary knows every place on earth. She has (even) been to BORneo!

b. People expected that John would win the election, followed by Bill, with Mary as a distant third. But then the election was won by MARY (out of all persons)!

c. John would distrust Albert SCHWEITzer!

(39.c) is an instance of what Fauconnier (1975) has called „quantificational superlatives". Assuming that Albert Schweitzer is a particularly trustworthy person, (39.c) expresses that John would distrust everyone.

The function of emphatic focus is to indicate that the proposition that is actually asserted is *prima facie* a particularly unlikely one with respect to the alternatives. This meaning component can be made explicit with particles like *even* or idiomatic constructions like *out of all persons*. Let us ssume that emphatic prosody indicates a particular type of assertion, emphatic assertion. It can be characerized to a certain degree as follows, where $p <_c q$ expresses that presupposition is less likely than presupposition q, given the information in the common ground c.

(40) **Emph.Assert**$(\langle B,F,A\rangle)(c) = c \cap B(F)$, iff

a) For all $F' \in A$: $c \cap B(F) <_c c \cap B(F')$

b) $c \cap B(F) <_c \cap\{c \cap B(F') | F' \in A\}$

Felicity condition (a) says that the assertion actually made, $c \cap B(F)$, is less likely in the current common ground c than any alternative assertion $c \cap B(F')$. In example (39.c), it must be considered less likely that John would distrust Albert Schweitzer than that he would distrust any other person. Condition (b) says that the assertion actually made is less likely in c than the conjunction of all the alternative assertions. In example, the common ground c must support the possibility that John would distrust all other persons but still does not distrust Albert Schweitzer. Only then the proposition that John would distrust Albert Schweitzer is a truly exceptional and unlikely one.

Note that the two conditions (43.a) and (b) are logically independent of each other. In particular, (a) does not entail (b), as the common ground c could contain the information that although Albert Schweitzer is the most trustworthy person, if someone distrusts every other person, then he distrusts Albert Schweitzer as well, and hence the left-hand side and the right-hand side of (b) would be equally likely. And (b) does not entail (a), as it might be that it is less probable that John distrusts Albert Schweitzer than that John distrusts all other persons together, but still there is one person (say, Mother Teresa) such that the propositions that John distrusts Albert Schweitzer and the proposition that John distrusts Mother Teresa are equally unlikely.

Now, a probability relation like $<_c$ is related to semantic strength $\subset$ in the following way: If p and q are comparable in their semantic strength (i.e. we have either $p \subseteq q$ or $q \subseteq p$), and furthermore $p <_c q$, then also $p \subset q$. That is, if p is less likely than q in c, then c allows for q-worlds that are not p-worlds, but not vice versa. Hence (43) amounts to the following condition for BFA-structures where the proposition expressed and its alternatives are related by semantic strength:

(41) If for all $F' \in A$: $c \cap B(F') \subseteq c \cap B(F)$ or $c \cap B(F) \subseteq c \cap B(F')$, then:

Emph.Assert($\langle B,F,A \rangle$)(c) = $c \cap B(F)$, provided that:

a) for all $F' \in A$: $c \cap B(F) \subset c \cap B(F')$

b) $c \cap B(F) \subset \cap\{c \cap B(F') | F' \in A\}$

The felicity condition (a) says that the proposition actually asserted, $c \cap B(F)$, must be stronger than every alternative proposition $c \cap B(F')$. And condition (b) says that that proposition must be stronger than the conjunction of all the alternative propositions.

If the alternatives are generated by a NPI the proposition expressed and its alternatives are indeed related by semantic strength, and hence emphatic assertion amounts to (41). It turns out that a sentence like (42.a) is indeed a good emphatic assertion, whereas a sentence like (42.b) is a bad emphatic assertion.

(42) a. Mary didn't get ANYthing.

b. *Mary got ANYthing.

Sentence (45.a) will yield the following BFA-structure:

(43) a. $\langle \lambda Q \lambda i \neg\exists y[Q_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)], \mathbf{thing}, \{P | P \subset \mathbf{thing} \wedge \neg\min(P)\} \rangle$

Applying **Scal.Assert** will give us a good result for common grounds c if the following conditions are satisfied:

(a) For all $P \in \mathbf{thing}$, $\neg\min(P)$: $\{i \in c | \neg\exists y[\mathbf{thing}_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)]\} \subset \{i \in c | \neg\exists y[P_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)]\}$, that is, the proposition that Mary didn't get a **thing** is not only as strong as, but stronger than any proposition that Mary didn't get some non-minor P, $P \subset$**thing**.

(b) $\{i \in c | \neg\exists y[\mathbf{thing}_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)]\} \subset \cup\{\{i \subset c | \neg\exists y[P_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)]\} | P \subset \mathbf{thing} \wedge \neg\min(P)\}$, that is, the proposition that Mary didn't get a **thing** is stronger than the conjunction of the propositions that Mary didn't get some non-minor P, $P \subset$**thing**. This is because the proposition that Mary didn't get a **thing** excludes that Mary even got a minor thing, whereas the conjunction of the alternative propositions does not exclude that.

Conditions (a) and (b) are satisfied for common grounds c that contain the information that it is prima facie less likely that Mary didn't get something including minor things than that Mary didn't get something excluding minor things. In other words, c must support the expectation that Mary got at least something minor, if not more. This is indeed the case for all common grounds in which a sentence like (45.a) is felicitous.

Sentence (45.b), on the other hand, will obviously lead to conditions that cannot be satisfied when emphatically asserted. In particular, condition (a) would amount to the requirement that for all $P \subset \mathbf{thing}$, $\neg\mathrm{min}(P)$: $\{i \in c | \exists y[\mathbf{thing}_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)]\} \subset \{i \in c | \exists y[P_i(y) \wedge \mathbf{get}_i(\mathbf{m},y)]\}$, that is, the proposition that Mary got a **thing** is stronger than the proposition that Mary got a P, where $P \subset \mathbf{thing}$. This is a clear contradiction.

4. The Distribution of Weak and Strong NPIs

One important question at this point is whether the semantics and pragmatics of assertions with weak and strong NPIs developed above captures the facts about their respective distribution. In particular, it should follow from the theory as developed so far that weak NPIs do not occur in emphatic assertions, and that strong NPIs do not occur in regular (scalar) assertions.

The first is a consequence of two facts: On the one hand, weak NPIs are exhaustive (cf. 24), that is, the union of their alternatives is equivalent to their meaning. On the other hand, emphatic assertions must be based on a meaning that is not only stronger than any alternative in particular, but also stronger than all the alternatives together (cf. 41.b). Hence weak NPIs are ruled out for emphatic assertions; in a sense, the meaning of a weak NPI is not „extreme" enough for a felicitous emphatic assertion. As strong NPIs are non-exhaustive (cf. 38), they are fine with emphatic NPIs.

The second consequence, that strong NPIs do not occur in regular assertions, can be motivated by assuming that the additional semantic condition for strong NPIs, namely, that their meaning is truly stronger than the union of their alternatives, is not exploited by regular assertions, and hence it is unmotivated to bring this condition into play in the first place. But it is unclear how to enforce this condition for non-exhaustive NPIs short of stipulating a general requirement for semantic compositions that they preserve the unique role of the foreground. The ultimate motivation may be Grice's maxim of relevance: If a speaker introduces NPIs with an „extreme" meaning, then the speaker should make appropriate use of this feature.

How does the present characterization of weak and strong NPIs and regular and emphatic assertion fit to Zwarts' observation, that strong NPIs are restricted to anti-additive contexts such as *no girl*, whereas weak NPIs can also occur in decreasing contexts such as *fewer than three girls*? In a previous attempt (Krifka 1995) I tried to argue that Zwarts' notion of anti-additivity should be replaced by the notion of strict decrease, where a function f is strictly decreasing iff if holds that whenever $X \subset Y$, then $f(Y) \subset f(X)$. However, I do not think that Zwarts' observation can be derived in this way (Jack Hoeksema, pers. communication).

I suspect that the following is behind Zwarts' observation: Emphatic assertions tend to be emphatic „across the board". That is, whenever there are expressions that are related to alternatives in an emphatic assertion, the meaning of the expressions has to be extreme with respect to the alternatives. An example of a good „across the board" emphatic assertion is the following:

(44) Bill is such a shrewd salesman; he would sell REFRIGERATORS to ESKIMOS.

Here, *refrigerators* introduces alternative sale items, and *Eskimos* introduces alternative customers; (46) is a good emphatic assertion because in (stereo-) typical contexts, to be able to sell refrigerators to Eskimos is considered to be less likely than, for example, to be able to sell walkmans to teenagers.

Now, it is plausible to assume that downward-entailing quantifiers come with alternatives, just like number words or upward-entailing quantifiers. For example, the alternatives to the meaning of *fewer than three* are the meanings of *fewer than four*, *five* etc. and *fewer than two* and *no*. The alternatives to the meaning of *no* are the meaning of *fewer than two*, of *fewer than three*, etc. Clearly, *no* is the extreme value with respect to this set of alternatives; *fewer than three* is just an intermediary value. Hence we should assume that *no* can occur easily in emphatic assertions with another strong NPI, whereas *fewer than three* should be resistant.

There is some evidence for this explanation of Zwarts' observation. For example, in cases like the following, the *no*-phrase is preferably read with strong, emphatic stress, just as the NPI itself:

(45) NO friend of mine lifted a FINGER / did ANYthing at ALL.

Furthermore, there are certain quantifiers that, while technically not anti-additive, seem to allow for strong NPIs. Examples are *hardly anyone* or *practically noone*, which have a meaning very similar to *few*, but seem to allow for strong NPIs:

(46) a. Hardly ANYONE lifted a FINGER to help me.

b. Practically NOONE lifted a FINGER to help me.

The reason why such examples are good may be that *anyone* and *noone* are extreme with respect to their alternatives, even though their extremity is somewhat toned down by the modifiers *hardly* or *practically*.

Furthermore, given that Zwarts' observations seem to be tendencies rather than strictly grammatical facts, we perhaps even do not want to rule out combinations like *fewer that three girls did anything at all* by fundamental principles.

5. Types of Polarity Items

In the previous sections we have discussed the general outline of the proposed theory with one particular example, *anything*. In this section I am going to discuss various types of polarity items: expressions of a general nature, operators that widen the applicability of a predicate, referentially non-specific expressions, and expressions that denote particularly small or large entities.

5.1. Expressions of General Nature

The NPI *anything* is an example of a NPI whose denotation, the property **thing**, is more general than any one of its alternatives. Other examples are noun phrases formed with the prefix or determiner any, such as *anybody* or *any girl*:

(47) a. *anybody*: $\langle B, \mathbf{person}, \{P| P\subset\mathbf{person}\}\rangle$ÿ20

b. *any girl*: $\langle B, \mathbf{girl}, \{P| P\subset\mathbf{girl}\}$

For example, *any girl* denotes the property **girl**, and has as alternatives a set of properties that are semantically stronger than **girl**. As with *anything*, I assume exhaustivity, that is, $\cup\{P| P\subset\mathbf{girl}\} = \mathbf{girl}$. There are also non-exhaustive variants, like *ANYbody (at ALL)* and *ANY girl (at ALL)* for which the alternative properties P are restricted to non-minor ones.

In addition to NPs based on *any*, we find a few idiomatic NPIs that also express concepts of a general nature, like sound or thing:

(48) a. John didn't hear a SOUND. (Alternatives: $\{P| P\subset\mathbf{sound} \wedge \neg\mathrm{min}(P)\}$)

b. John didn't eat a THING. (Alternatives: $\{P| P\subset\mathbf{edible.thing} \wedge \neg\mathrm{min}(P)\}$)

The meaning of *sound* includes any acoustical event, and the meaning of *thing* includes every context-relevant object, more narrowly specified by the sortal restrictions imposed by the verbal predicate. This general meaning is the source for their idiomatization as NPIs, which essentially consisted in getting conventionally associated with alternatives. It seems that these NPIs are obligatorily focussed, that is, in emphatic assertions, and hence are strong, or non-exhaustive, NPIs. The expressions *a sound* and *a thing* in their idiomatic uses have the same meaning as *any sound at all* and *anything at all*.

5.2. Expressions that Relax Criteria of Applicability

Another type of NPIs are expressions like *much of a*, *at all* or *in the least* in examples like the following:

(49) a. Mary isn't much of a clarinetist.

b. *Mary is much of a clarinetist.

(50) a. John isn't tired at all.

b. *John is tired at all.

These expressions induce a most liberal interpretation of the expressions in its scope. This may be with respect to how strict vague predicates are interpreted, as in (54), or with respect to the reasons or evidence for the application of a predicate, as in the following example:

(51) a. Mary isn't pregnant at all.

b. *Mary is pregnant at all.

I would like to propose that a phrase like *tired at all* has the structure of a NPI, with the meaning of *tired* interpreted in the most liberal way or requiring the least evidence, and a set of alternatives that consists of the meaning of *tired* interpreted in stricter ways or requiring greater evidence.

One way of implementing this idea is to interpret constants with respect to different „precision standards“, something that has been proposed for degree adjectives by Lewis (1970), or „standards of evidence“. In particular, we may assume that indices contain a component that specifies more or less strict ways of interpreting a predicate or more or less strict evidence for applying a predicate to a particular individual, and that indices are ordered according to the strictness of interpretation and the evidential support they require (cf. Landman 1991). Let $\leq_s$ such an order, where $i \leq_s j$ means that standard i is at least as strict as standard j. The relation $\leq_s$ is defined as follows:

(52) $i \leq_s j$ iff
a) i and j differ at most in their precision standard,
b) for all constants α, $\alpha_i \subseteq \alpha_j$

Clause (b) allows for the extension of constants to increase with decreasing precision standards. For example, if John is not tired at i, he may count as tired at the less strict standard j. The expression *at all* as a property modifier then has the following meaning:

(53) **at.all** (as a predicate modifier):
$\lambda P.\langle \lambda Q.Q,\ \lambda i \lambda x \exists j[i \leq_s j \wedge P_j(x)],\ \{\lambda i \lambda x[i=j \wedge P_i(x)] | \exists k[j \leq_s k \wedge P_j \subset P_k]\}\rangle$
abbreviated: $\lambda P.\langle \lambda Q.Q,\ \mathbf{at.all},\ \mathbf{at.all}^A\rangle$

We get the meaning of *tired at all* by applying (53) to the property **tired**:

(54) $\langle \lambda Q.Q,\ \mathbf{at.all}(\mathbf{tired}),\ \{X(\mathbf{tired}) | X \in \mathbf{at.all}^A\}\rangle$
$= \langle \lambda Q.Q,\ \lambda i \lambda x \exists j[i \leq_s j \wedge \mathbf{tired}_j(x)],$
$\{\lambda i \lambda x[i=j \wedge \mathbf{tired}_i(x)] | \exists k[j \leq_s k \wedge \mathbf{tired}_j \subset \mathbf{tired}_k]\}\rangle,$
abbreviated: $\langle \lambda Q.Q,\ \mathbf{tired.at.all},\ \mathbf{tired.at.all}^A\rangle$

The foreground of *tired at all* is the property that holds of all the individuals that are tired under some (possibly weaker) precision standard j. The alternatives are the meanings of *tired* under some precision standard that is not the minimal one for *tired*. Under this construction, the foreground is weaker than every alternative. In other words, for every P, $P \in \mathbf{tired.at.all}^A$, it holds that $P \subset \mathbf{tired.at.all}$. Our two examples (50.a,b) then are interpreted in the following way:

(50’) a. $\langle \lambda Q \lambda i[\neg Q_i(\mathrm{JOHN})],\ \mathbf{tired.at.all},\ \mathbf{tired.at.all}^A\rangle$

b. $\langle \lambda Q \lambda i[Q_i(\mathrm{JOHN})],\ \mathbf{tired.at.all},\ \mathbf{tired.at.all}^A\rangle$

Phrases modified by *at all* clearly are strong NPIs, as they require emphatic stress. This is captured by our reconstruction by the fact that they are non-exhaustive. For example, the foreground of *tired at all* is the property of being tired to some, including a minimal, degree, whereas the alternatives are properties of being tired to some non-minimal degree. Clearly, there are entities that are tired to a minimal degree but not tired to a non-minimal degree, and hence we have $\cup$**tired.at.all**$^{A} \subset$ **tired.at.all**.[10]

It is obvious from several examples that we have discussed so far that *any* and *at all* can be combined, as in *Mary didn't get any present at all*, yielding a strong NPI. It is possible to analyze expressions like *any present at all* compositionally: *any present* is represented by $\langle$B, **present**, {P| P$\subset$**present**}$\rangle$, and *at all* induces a widening of the precision standard for **present**. Hence we get the following representation, which also illustrated the interpretation of *at all* when it is applied to a BFA structure:

(55) *any present*: $\langle B, \mathbf{present}, \{P|\ P\subset \mathbf{present}\}\rangle$

at all: $\lambda\langle B,F,A\rangle[\lambda Q.B(Q), \mathbf{at.all}(F), \{X(Y)|\ X\in \mathbf{at.all}^{A} \wedge Y\in A\}\rangle$

any present at all: $\langle B, \lambda i\lambda x\exists j[i\leq_s j \wedge \mathbf{present}_j(x)],$

$\{\lambda i\lambda x[i=j \wedge P_i(x)]|\ \exists k,P[P\subset\mathbf{thing} \wedge j\leq_s k \wedge P_j\subset P_k]\}\rangle$

The foreground part is the property of being a present at some weaker precision standard j. The alternatives are properties P' that are subproperties of **thing** and that are interpreted at some non-minimal precision standard. This should substitute the preliminary representation that was using a special predicate „min“.

5.3. Referentially Non-Specific Expressions

Let us turn to another type of NPIs, which can be illustrated by ever:

(56) a. It is not the case that Mary has ever been to China.

b. *Mary has ever been to China.

Heim (1987) has analyzed *ever* as meaning ‘at least once’, and as introducing alternatives like ‘at least n times’, for n>1. But what seems to be relevant for an example like (56) is not the number of events (which would be focused on by stressed *a single time*), but that the speaker does not refer to any specific time or event. Hence I assume that the meaning of *ever* suppresses reference to specific times.

Supporting evidence for this assumption comes from data like the following, in which a temporal adverbial specifies a reference time.

(57) When I left home yesterday, I didn't (*ever) close the windows.

If the function of a specific *when*-clause is to introduce a reference time that is to be taken up by the main clause, then (57) is bad because *ever* prevents the main clause from doing its expected job.

A full reconstruction of the semantics of *ever* requires a framework which incorporates quantification over events and reference times, such as Partee (1984). Here I just want to illustrate the principal ingredients with the means at hand, reference to indices, which may include times as one component. The analysis then can be recast in one's favourite framework of quantification over events or situations.

Assume that sentences have a reference time parameter t that normally is fixed either by temporal adverbials or by anaphoric reference to some salient reference time. Assume that this reference time parameter is part of the index i, which is conceived of as a pair of a world and a time interval $\langle w,t\rangle$. The relation AT should hold between a time interval, a proposition, and an index such that AT(t',p,i) is true iff t' is the time of an event that satisfies the proposition p at i. For example, **at**(t', **I.left.home.yesterday**, $\langle w,t\rangle$) holds iff t' is an interval of me leaving home yesterday, interpreted with respect to the world w and the time t. Then the *when*-clause in (57) can be represented as in (58), and the whole sentence (57) as in (59):

(59) . $\lambda p\lambda\langle w,t\rangle[p(\langle w,\iota t'[\mathbf{at}(t',\mathbf{I.left.home.yesterday},\langle w,t\rangle)]\rangle)]$

(60) $\lambda\langle w,t\rangle[\neg\mathbf{I.close.the.window}(\langle w,\iota t'[\mathbf{at}(t',\mathbf{I.left.home.yesterday},$
$\langle w,t\rangle)]\rangle)]$

(60) is a proposition that maps $\langle w,t\rangle$ to truth if it is not the case that I closed the window at t', the time at which I left home in w at the day preceding t. Now, the function of *ever* is to existentially bind the time parameter. The alternatives of a sentence containing *ever* are propositions for which the time parameter is set to some value or other. Hence we have the following interpretation, where *ever* is treated as a proposition modifier for simplicity.

(60) $\langle\lambda X.X,\ \lambda p\lambda\langle w,t\rangle\exists t'[p(\langle w,t'\rangle)],\ \{\lambda p\lambda\langle w,t\rangle[p(\langle w,t'\rangle)]|\ t'\in T\}\rangle$
where T is the set of contextually relevant times.

According to this representation the bad version of (57) is out because the adverbial clause fails to specify a time for the main clause. That is, when the foreground of (60) is applied to the proposition expressed by *I didn't close the window*, we get a proposition $\lambda\langle w,t\rangle\exists t'[\neg\mathbf{I.close.window}(\langle w,t'\rangle)]$ whose relevant temporal parameter t' cannot be fixed by operators like the adverbial clause in (58).

Let us see what our semantic and pragmatic rules tell us about sentences involving *ever*. Our two examples in (60) get the following representation:

(61) a. $\langle\lambda O[O(\lambda i[\neg\mathbf{Mary.go.to.China}(i)])],$
$\lambda p\lambda\langle w,t\rangle\exists t'[p(\langle w,t'\rangle)],\ \{\lambda p\lambda\langle w,t\rangle[p(\langle w,t'\rangle)]|\ t'\in T\}\rangle$

b. $\langle\lambda O[O(\lambda i[\mathbf{Mary.go.to.China}(i)])],$
$\lambda p\lambda\langle w,t\rangle\exists t'[p(\langle w,t'\rangle)],\ \{\lambda p\lambda\langle w,t\rangle[p(\langle w,t'\rangle)]|\ t'\in T\}\rangle$

In (61.a), the asserted proposition, $\lambda\langle w,t\rangle\neg\exists t'[\mathbf{Mary.go.to.China}(\langle w,t'\rangle)]$, is at least as strong as any alternative assertion $\lambda\langle w,t\rangle[\neg\mathbf{Mary.go.to.China}(\langle w,t'\rangle)]$, for every $t'\in T$. Informally, the assertion that Mary didn't got to China at some time or other implies that Mary didn't go to China at some specific time t'. This is the configuration for good assertions. In (61.b), the asserted proposition is at least as weak as any alternative assertions, which explains why it is bad.

NPIs based on referentially non-specific expressions are exhaustive under the natural assumption that the set of alternative reference times T contains all the times the existential quantifier in the foreground can range over. We then have $\{\langle w,t\rangle | \exists t'[p(\langle w,t'\rangle)]\} = \cup\{\{\langle w,t\rangle | [p(\langle w,t'\rangle)]\} | t' \in T\}$. Hence such NPIs are predicted to be weak, which is indeed the case.

5.4. *Expressions that Refer to Minimal or Maximal Entities*

Another type of NPIs are predicates that refer to very small entities of a certain sort.

(62) a. John didn't drink a drop (of alcohol) for two days.
b. Mary didn't utter {a word / a syllable}.
c. John doesn't have a red cent.

Take *a drop* as an example. In its NPI use it applies to minimal liquid quantities[11], and its alternative predicates apply to bigger liquid quantities. We can make this more precise as follows. Assume that $\subset$ expresses the proper part relation; $x \subset_i y$ says that x is a proper part of y at index i.

(63) *a drop*: $\langle \lambda Q.Q, \mathbf{drop}, \mathbf{drop}^A \rangle$, where

$\mathbf{drop} = \lambda i.\{x | \mathbf{liquid}_i(x) \wedge \neg\exists y[y \subset_i x]\}$, and

$\mathbf{drop}^A$ is a set that satisfies the following requirements:

a) $\forall i \forall x[\mathbf{liquid}_i(x) \rightarrow \exists P[P \in \mathbf{drop}^A \wedge P_i(x)]]$

b) $\forall i \forall P \forall x[P \in \mathbf{drop}^A \wedge P_i(x) \rightarrow \mathbf{liquid}_i(x)]$

c) $\forall i \forall P \forall P'[P \in \mathbf{drop}^A \wedge P' \in \mathbf{drop}^A \wedge P \neq P' \rightarrow \neg\exists x[P_i(x) \wedge P'_i(x)]]$

In prose, **drop** is a property that refers to all minimal quantities of liquid, that is, to all quantities of liquid x that do not contain proper parts. The set of alternatives, $\mathbf{drop}^A$, is such that (a) for each index i, if x is a quantity of liquid, then there is some property P that applies to x, (b) for each index i and property P, P applies only to quantities of liquid, and (c) the properties P are disjoint. Conditions (a)-(c) are necessary conditions that may be further refined, for example by requiring that each alternative property applies to quantities of liquid of a certain size. I am aware that conditions (a) to (c) do not define a unique $\mathbf{drop}^A$, but I will not be more specific here as any set of properties that satisfies them will do for our purposes.

Other NPIs of this type can be analyzed in a similar fashion. For example, *a word* or *an iota* denotes minimal utterances, *a red cent* denotes minimal amounts of money, *lift a finger* denotes minimal amounts of labor, and *bat an eye* applies to the least reactions to threatening events. It is obvious that these expressions have to be understood in their non-literal meaning: They are idiomatic expressions that denote minimal elements of certain ontological sorts.

Now, observe that NPIs like *a drop* and their ilk are not directly based on informativity under the reconstruction given above. However, they lead to alternative assertions based on informativity under a certain plausible assumption (cf. also Fauconnier 1980). It is perhaps best to discuss this using an example:

(64) a. *Mary drank a drop.

$\langle \lambda Q\{i|\exists y[Q_i(y) \wedge \mathbf{drank}_i(\mathbf{m},y)]\}, \mathbf{drop}, \mathbf{drop}^A\rangle$

b. Mary didn't drink a drop.

$\langle \lambda Q\{i|\neg\exists y[Q_i(y) \wedge \mathbf{drank}_i(\mathbf{m},y)]\}, \mathbf{drop}, \mathbf{drop}^A\rangle$

We want to derive that (66.a) is bad as an assertion, whereas (66.b) is good. We can do so under the plausible assumption that if someone drinks something, he drinks every part of it. Let us call this principle, in general, „involvement of parts“:

(65) $\forall i \forall x \forall y \forall z[\mathbf{drink}_i(x,y) \wedge z \subset_i y \rightarrow \mathbf{drink}_i(x,z)]$

A corollary to (67) is: If someone drinks some quantity of liquid, he also drinks minimal quantities, as every quantity of liquid will contain minimal quantities.

(66) $\forall i \forall x \forall y[\mathbf{drink}_i(x,y) \wedge \mathbf{liquid}_i(y) \rightarrow \exists z[\mathbf{drop}_i(z) \wedge \mathbf{drink}_i(x,z)]]$

A second principle is that the predicate **drop** applies to liquid quantities of an idealized small size. We can capture this by requiring of natural common grounds c that the proposition that someone drank just a minimal quantity of liquid at not more should always be less probable than that he or she drank a more substantial quantitiy of liquid. Let us call this the „principle of extremity“:

(67) For all natural common grounds c and all x, y:

$\{i|\ \mathbf{drink}_i(x,y) \wedge \mathbf{drop}_i(y) \rightarrow \neg\exists z[y \subset_i z \wedge \mathbf{drink}(x,z)]\} <_c$

$\{i|\ \mathbf{drink}_i(x,y) \wedge \mathbf{drop}_i(y) \rightarrow \exists z[y \subset_i z \wedge \mathbf{drink}(x,z)]\}$

Let us come back to examples (64.a,b) in the light of these principles. First note that the NPIs in question are all strong; they bear heavy stress and can easily be combined with *even*. Hence we should assume emphatic assertion. In (64.a), the proposition asserted with respect to the input common ground c, $\{i \in c|\exists y[\mathbf{drop}_i(y) \wedge \mathbf{drank}_i(\mathbf{m},y)]\}$, is at most as strong as any alternative assertion $\{i \in c|\exists y[P_i(y) \wedge \mathbf{drank}_i(\mathbf{m},y)]\}$, $P \in \mathbf{drop}^A$, according to involvement of parts (65), and in fact weaker if c is a natural common ground according to extremity (67). This directly contradicts condition (41.a) for emphatic assertions. In (66.b), the proposition asserted with respect to c, $\{i \in c|\ \neg\exists y[\mathbf{drop}_i(y) \wedge \mathbf{drank}_i(\mathbf{m},y)]\}$, is truly stronger than any alternative assertion $\{i \in c|\neg\exists y[P_i(y) \wedge \mathbf{drank}_i(\mathbf{m},y)]\}$, $P \in \mathbf{drop}^A$ for every natural common ground c due to extremity, which abides by condition (41.a).

The principle of extremity has an interesting consequence. Without it it should be possible to use a sentence like *Mary drank a drop* to express that Mary drank only minimal amount. This may even be possible in ironic or hyperbolic uses. The principle of extremity, however, excludes that, as it would hold in the output common ground that the probability that Mary just drank a drop is 1, wheres the the probability that Mary drank more than a drop is 0.

Other NPIs of this type, like *lift a finger* or *a red cent*, can be explained in a similar way. Interestingly, there are a few NPIs that are based on predicates that denote „large“ entities:[12]

(68) a. Wild horses couldn't drag me in there.

b. We will not know the truth {in weeks / in a million years}.

The basic reasoning is quite similar to the former case. For example, *in weeks* and *in a million years* refers to a time that is maximally distant in the future with respect to a given context. We assume a general rule that, if a person knows something at a time t, then he knows it at any time t' later than t. Then the claim that we don't know it at a time that is maximally distant in the future is stronger than the claim that we don't know it at some other time. In addition, the extremity principle says in the case at hand that it is less likely that we will know the truth only at the most distant future time, than that we know the truth already at some earlier time. This is the setting that results in good emphatic assertions.

It should be immediately obvious that NPIs based on small or large entities are not exhaustive. Take the case *a drop*; if **drop** applies just to minimal liquid quantities and all the altnernatives in $\mathbf{drop}^A$ apply to bigger liquid quantities, then we have not only $\mathbf{drop} \neq \cup\mathbf{drop}^A$, but even $\mathbf{drop} \cap \cup\mathbf{drop}^A = \emptyset$. And when we take larger expressions that contain *a drop*, like *drink a drop*, then we find, due to involvement of parts and extremity, that for every natural context c, $\lambda i{\in}c\lambda x\exists y[\mathbf{drop}_i \wedge \mathbf{drink}_i(x,y)] \subset \cup\{\lambda i{\in}c\lambda x\exists y[P_i \wedge \mathbf{drink}_i(x,y)]|\ P{\in}\mathbf{drop}^A\}$, as those worlds in which someone just drank a drop are considered most unlikely.

It should be noted that all the polarity items discussed in this section also have a literal meaning in which they do not act as a polarity item. A sentence like *He drank* (or *did not drink*) *a drop of alcohol* could mean: *He drank* (or: *did not drink*) *a quantity of alcohol falling in a spheroical mass* (one dictionary's definition of *drop*). The polarity use can be seen as a case of grammatization, the semantic change from a rather specific meaning to a much wider meaning that is related to semantic sorts.

5.5. Positive Polarity Items

The theory developed above can be applied to positive polarity items (PPIs), as in thefollowing case:

(69) a. John has TONS of money.

b. *John doesn't have tons of money.
[o.k. as a denial of (a) or with contrastive focus on *tons*]

The expression *tons of* forms PPIs. For example, *tons of money* applies to maximal amounts of money, i.e. amounts of money that are higher than some very high threshold value, and its alternatives are properties that apply to smaller amounts of money. We can assume involvement of parts: If John owns x, John also owns the parts of x. Furthermore, we can assume extremity: For every natural context c it holds: that someone has less than a maximal amount of money is less likely than that someone has a maximal amount of money. Then the proposition that John owns a maximal amount of money is stronger than any proposition that John owns some other amount. According to the by now familiar scheme, this makes (69.a) a good assertion. On the other hand, the proposition that John doesn't own a maximal amount of money is weaker than the proposition that John doesn't own some other amount, and hence (69.b) is a bad assertion.

There are also non-idiomatic PPIs. Let us discuss the PPI rather, or pretty, as a predicate modifier:

(70) a. John is rather/pretty tired.

b. *John isn't rather/pretty tired.

Rather in this use[13] can be seen as the counterpart of the NPI *at all*, as it introduces alternatives that are interpreted more liberally. But contrary to *at all*, which may quantify over degrees of evidence, *rather* seems to quantify only over interpretation standards for vague predicates (cf. **rather pregnant*). I propose the following meaning and alternatives for *rather tired*:

(71) a. **rather.tired**: $\lambda i \lambda x[\mathbf{tired}_i(x)]$

b. **rather.tired**^{A}: $\{\lambda i \lambda x[i \leq_s j \wedge \mathbf{tired}_j(x)] | j \in I\}$

The meaning of *rather tired* is the meaning of *tired*, at some given precision standard that comes with the index i, and the alternatives are interpretations of *tired* at weaker precision standards. Given this analysis, we can derive the distribution of *rather tired* in the usual way. In particular we always have $\mathbf{rather.tired} \subseteq F'$, for all F', $F' \in \mathbf{rather.tired}^A$. We also have $\mathbf{rather.tired} = \cup \mathbf{rather.tired}^A$, which means that *rather tired* is a exhaustive, and hence a weak PPI.

The reader might wonder why I do not include a treatment of NPs like *something* that often are considered positive polarity items. As I explained with example (34) above, I think that *something* is not a polarity item and does not introduce any alternatives.

5.6. Presuppositional Polarity Items

As a last class of polarity items let me mention those that introduce the required logical conditions through a special presupposition. Take, for example, the PPI *already*. For simplicity, let us concentrate on the temporal, non-focussing use illustrated in the following example:

(72) a. Mary is already here.

b. *Mary isn't already here.

The adverb *already* expresses that the sentence in its scope is true at the time t of its index, and introduces alternative times t' later than t such that the sentence is true at t'. For example, (72.a) asserts the proposition (i) in contrast to the alternatives (ii), where t<t' means that t' is later than t. Furthermore, *already* comes with the presupposition (iii) that the sentence changes its truth value from false to true and stays true within the contextually relevant time.[14]

(73) a. (i) $\lambda\langle w,t\rangle.\mathbf{Mary.is.here}(\langle w,t\rangle)$

(ii) $\{\lambda\langle w,t\rangle| \mathbf{Mary.is.here}(\langle w,t'\rangle)]| t<t'\}$

(iii) Presupposition: $\lambda\langle w,t\rangle.\exists t'\forall t''[[t''\langle t' \rightarrow \neg\mathbf{Mary.is.here}(\langle w,t''\rangle)]$
$\wedge [t'<t'' \rightarrow \mathbf{Mary.is.here}(\langle w,t''\rangle)]]$

Under the presupposition (iii) the proposition (i) is indeed stronger than any alternative: Whenever $\mathbf{Mary.is.here}(\langle w,t\rangle)$ is true, then $\mathbf{Mary.is.here}(\langle w,t'\rangle)$ will be true for times t' after t, but not vice versa.

The PPI *still* is similar to *already* except for the temporal orientation of its alternatives and its presupposition. And the NPI *yet* resembles *already* except for the temporal orientation of its alternatives.

6. The Locus of Exploitation of Polarity Items

6.1. „Doubly-licensed“ Polarity Items

Under the semantico-pragmatic account of polarity items we would expect that polarity items under more than one licensing operator show a flip-flop behavior. This is indeed attested in certain cases. Baker (1970) pointed it out for PPIs with examples of the following kind:[15]

(74) a. I would rather be in Montpellier.

b. ??I wouldn't rather be in Montpellier.

c. There isn't anyone in the camp who wouldn't rather be in Montpellier.

Sentence (75.b) is acceptable only if the concept of „would rather be in Montpellier“ has been mentioned before; typically, either *I* or *wouldn't* are stressed in these cases. — Schmerling (1971) showed that we find a similar „flip-flop“ behavior with NPIs:

(75) a. *There was someone who did a thing to help.

b. There was no one who did a thing to help.

c. *There was no one who didn't do a thing to help.

These grammaticality judgements can be immediately explanined from the semantics of licensers, here negation, as two negations cancel each other ($\neg\neg\phi = \phi$).

However, there are cases where an NPI occurs in the scope of two licensing operators, which seems to be a true paradox for any semantic theory of polarity items. Hoeksema (1986) discusses cases of NPIs in the protasis of conditionals like (77), and Dowty (1994) presents cases of NPIs in the scope of downward-entailing adverbial quantifiers (cf. 78):

(76) a. If he knows anything about logic, he will know Modus Ponens.

b. If he doesn't know anything about logic, he will (still) know Modus Ponens.

(77) a. She very rarely eats anything at all for lunch.

b. She very rarely doesn't eat anything at all for lunch.

Ladusaw (1979) was aware of these facts: The implementation of his theory requires that an NPI be licensed by one downward-entailing operator; once licensed, it will stay licensed. Dowty (1994) suggests a distinction between semantic licensing based on downward-entailingness, and syntactic licensing that suppresses the flip-flop behavior of semantic licensing.

The solutions that have been presented for doubly-licensed NPIs are problematic for the semantico-pragmatic account of polarity items as they work with various principles that are extraneous to the idea that polarity items are used to express relatively „strong“ propositions. In this section I will argue that we can treat these phenomena within a semantic theory if we allow for a more flexible way of how the semantic contribution of polarity items is exploited.

6.2. Flexible Exploitation of Polarity Items

I would like to propose that the semantic contribution of a polarity item can be exploited at various levels of a complex semantic expression, not just at the uppermost level of the sentence. Independent evidence for this comes from cases like the following one:

(78) The student who had not read anything gave improvised answers.

Following the theory developed so far, (79) would be analyzed as follows: The NPI *anything* introduces alternatives in the usual way. These alternatives are projected in semantic compositions, and the negation in the relative clause reverses the specificity ordering. The assertion operator then makes use of the resulting alternatives:

(79) $\mathbf{Assert}(\langle \lambda Q \lambda i.\mathbf{gave.improvised.answers}_i(\iota x[\mathbf{student}_i(x) \wedge \neg\exists y[\mathbf{read}_i(x,y) \wedge Q_i(y)]]), \mathbf{thing}, \{P |\, P \subset \mathbf{thing}\}\rangle$

The problem is that the definite NP interrupts the semantic specificity relation between the foreground **thing** and its alternatives and the resulting propositions. For example, if John is the student who had not read anything, then replacing **thing** by some alternative P, P⊂**thing** will either give us the same proposition, or it will result in a presupposition failure (if there is another student who did read something but not P). Hence (79') cannot be an adequate representation of (79).

Obviously the NPI in (79) is licensed locally in its clause. Assuming that the alternatives introduced by polarity items are always exploited by illocutionary operators we have to assume that such operators can occur in embedded sentences:

(80) **Assert**[The student [**Assert** who had not read anything] gave improvised answers]

It is the downstairs **Assert** operator that makes use of the alternatives introduced by the NPI. In doing so this operator will neutralize these alternatives, making them unavailable for the upstairs **Assert** operator.

In order to implement this idea we must develop a framework in which illocutionary operators are part of the semantic recursion. This can be done when we assume that semantic representations, in general, are dynamic, that is, functions from input information states to output states. For ease of exposition I will not give recursive dynamic rules for subclausal expressions; see Krifka (1993) for how this can be done for BFA structures. We may define the dynamic version of a proposition p from its static version p' as follows: $p = \lambda c[c \cap p']$.

The rules for assertion will get a slightly different format. First, simple assertion is functional application, perhaps with the additional requirement of assertability, i.e. that the asserted sentence is compatible with and not already entailed by the input state:

(81) $\mathbf{Assert}(p) = \lambda c\iota c'[c'=p(c)\ (\wedge\ c'\neq c' \wedge c\neq\emptyset)]$

As before I will suppress the part in parentheses. How should we define assertion for BFA structures? Following our earlier analysis (29), we may suggest the following:

(82) $\mathbf{Assert}(\langle B,F,A\rangle) =$

$$\lambda c\iota c'[c'=B(F)(c) \wedge \forall F'\in A[B(F)(c) \neq B(F')(c) \rightarrow \text{Speaker has reasons not to propose } c'=B(F')(c)]]$$

In case the alternative propositions are related to B(F) by informativity we may assume a special operator **Scal.Assert**, as in (31). In the present framework this operator can be rendered as in (83): The input state c is changed to one in which B(F) is true and all alternative propositions B(F')(c) that are stronger than B(F)(c) are false:

(83) $\mathbf{Scal.Assert}(\langle B,F,A\rangle) = \lambda c[B(F)(c) - \cup\{B(F')(c) | F'\in A \wedge B(F')(c)\subset B(F)(c)\}]$

If the BFA structure is generated by a polarity item, then B(F')(c) is never stronger than B(F)(c) in the felicitous case, hence the $\cup$-term reduces to $\emptyset$, and we get B(F)(c) as output. In the infelicitous case the $\cup$-term either equals B(F)(c), which will yield $\emptyset$ as output, or it covers all non-extreme cases, which will leave only an extreme output state.

Our example (79), assuming simple polarity assertion, can be analyzed as follows:

(84) $\mathbf{Assert}(\lambda c[\{i\in c|\ \mathbf{gave.improvised.answers}_i(\iota x[\mathbf{student}_i(x) \wedge$

$$\mathbf{Scal.Assert}(\langle\lambda Q\lambda c.\{i\in c|\ \neg\exists y[\mathbf{read}_i(x,y) \wedge Q_i(y)]\},\ \mathbf{thing},\ \{P|P\subset\mathbf{thing}\}\rangle)(c)(i)])\}])$$

Notice that the contribution of the NPI is evaluated by **Scal.Assert**, which returns a function from information states c to information states c' such that in all worlds i of c', x (the student) hasn't read anything.

I think that the paradoxes of double licensing can be explained in a similar way. For example, a case like (78.b) may be analyzed as follows. *Rarely* expresses a quantification over lunch-situations s in which Mary takes part. We may analyze it as a relation $\lambda X\lambda Y[\#(X\cap Y)<n]$, where n is a small threshold value. **Eat** is a three-place predicate that relates an eater, an object that is eaten, and a situation s.

(85) Mary rarely doesn't eat anything for lunch.

$$\textbf{Assert}(\lambda c\{i\in c|\ \textbf{rarely}(\{s|\ \textbf{lunch}_i(s)\},$$
$$\{s|\ \textbf{Scal.Assert}(\langle\lambda Q\lambda c\{i\in c|\exists x[\textbf{eat}_i(\textbf{m},x,s)\wedge Q_i(x)]\},$$
$$\textbf{thing},\ \{P|\ P\subset\textbf{thing}\}\rangle)(c)(i)\})\})$$

Notice that the occurrence of the NPI is licensed locally, by **Scal.Assert**. The upstairs illocutionary operator is a simple assertion that does not relate to the alternatives introduced by the NPI.

The contrasting case (77.a), of course, is one in which there is no embedded illocutionary operator, and the alternatives introduced by the NPI do affect the illocutionary operator of the sentence:

(86) Mary rarely eats anything for lunch.

$$\textbf{Scal.Assert}(\langle\lambda Q\lambda c\{i\in c|\ \textbf{rarely}(\{s|\ \textbf{lunch}_i(s)\},$$
$$\{s|\ \exists x[\textbf{eat}_i(\textbf{m},x,s)\wedge Q_i(x)]\})\},\textbf{thing},\ \{P|\ P\subset\textbf{thing}\}\rangle)$$

(86) is a good assertion because **rarely** allows for downward-entailing inferences in its second argument: If Mary rarely eats vegetables, then Mary rarely eats carrots.

The examples involving a conditional, (76.a,b) can be treated in a similar way. First, it can be shown that (76.a) is a good scalar assertion under the standard analysis of (indicative) conditionals in dynamic interpretation (cf. Stalnaker 1975):

(87) $\textbf{if}(p,q) = \lambda c[[c\text{–}p(c)] \cup q(p(c))]$

That is, *if p then q* changes an input c to a c' that does not allow for p-worlds that are not q worlds. Notice that, due to the set subtraction in the first term, we have that if $p'(c)\subseteq p(c)$ then $\textbf{if}(p,q)(c)\subseteq\textbf{if}(p',q)(c)$. This is the reason why NPIs can occur in the protasis of conditionals. Our example (77.a) will be analyzed in the following way, were **jk**(Q) should represent „John knows about Q“, and **jkmp** stands for „John knows Modus Ponens“:

(88) a. If John knows anything about logic, he knows Modus Ponens.

$$\textbf{Scal.Assert}(\langle\lambda Q[\textbf{if}(\textbf{jk}(Q),\textbf{jkmp})],\ \textbf{logic},\ \{P|\ P\subset\textbf{logic}\}\rangle)$$
$$= \lambda c[\textbf{if}(\textbf{jk}(\textbf{logic}),\textbf{jkmp})(c) -\ \cup\{\textbf{if}(\textbf{jk}(P),\textbf{jkmp}(c)|\ P\subset\textbf{logic}\wedge$$
$$\textbf{if}(\textbf{jk}(P),\textbf{jkmp})(c)\subset\textbf{if}(\textbf{jk}(\textbf{logic}),\ \textbf{jkmp})(c)\}]$$

The input common ground c is first restricted to the set of worlds in which it holds that if John knows something about logic then he knows Modus Ponens. From this set the union of all those specifications of c is subtracted for which it holds that the proposition „if John knows something about logic then he knows Modus Ponens“ is stronger at c than the proposition „if John knows something about P then he knows Modus Ponens“, for $P\subset$**logic**. Due to the interpretation of conditionals (87) there is no such information state, hence that union is the empty set.

Example (76.b) can be explained by assuming local exploitation of the polarity structure. Rather informally we can assume the following analysis for (76.b):

(89) b. **Assert**[**if**(**Scal.Assert**[John doesn't know anything about logic],
he will not know Modus Ponens.]

So much about local exploitation of polarity items.[16] One obvious question at this point is, of course, where local exploitation can be applied.

6.3. Where can Polarity Items be Exploited?

We may assume that polarity items can be exploited at every clausal level, as examples like (76.b) and (78) show. However, notice that (75.c) should then be grammatical, as the NPI would be licensed in its local clause. It seems to me that the grammaticality judgements for these sentences are indeed questionable. They may be due to the fact that sentences (75.a), (b) and (c) are presented together and a certain interpretation — the one with a single, wide-scope illocutionary operator — is kept constant for every sentence.

Examples like (77.b) show that alternatives can be exploited even at a sub-clausal level. It may matter that the locus of exploitation in the semantic interpretation is the nuclear scope of a quantifier.

Another, related question is: What forces the assumption of operators that make use of alternatives? I think that the general principle is that a sentence must end up as being pragmatically well-formed. Consider the following cases:

(90) a. **Scal.Assert**[Mary rarely eats anything for lunch]

b. **Assert**[Mary rarely **Scal.Assert**[doesn't eat anything] for lunch]

c. ***Scal.Assert**[Mary rarely doesn't eat anything for lunch]

d. ***Scal.Assert**[Mary rarely **Scal.Assert**[doesn't eat anything] for lunch]]

As we have seen, (90.a,b) are pragmatically well-formed. (90.c) is bad because there is no information state that would satisfy the requirements of **Scal.Assert**. And (86.d) is bad as the NPI alternatives, so to speak, are already „used up" by the first **Scal.Assert** operator.

7. Locality Restrictions

One type of phenomenon that seems to argue for a syntactic treatment of polarity items are the various locality restrictions that have been observed, especially by Linebarger. In this section I will show that a semantic treatment of locality phenomena seems feasible as well.

7.1. Projection Failure

One kind of phenomenon that has been described as showing syntactic island effects for NPIs can be traced back to the failure of certain semantic constructions to project BFA structures properly. Take the contrast between the following sentences which shows that definite NPs, but not indefinite (non-specific) NPs impose restrictions for licensing of NPIs:

(91) a. Mary never goes out with men who have any problems.

b. *Mary never goes out with the man who has any problems.

This contrast can be explained by the current theory because the definite NP in (91.b) does not project the BFA-structure introduced by the NPI, whereas the nonspecific NP in (a) does. For (b) we would get the following BFA-structure:

(92) $$\langle\lambda Q\lambda i\neg[\mathbf{go.out.with}_i(\mathbf{m},\iota x\exists y[\mathbf{man}_i(x) \wedge Q_i(y) \wedge \mathbf{have}_i(x,y)])],\ \mathbf{problem},\ \{P|\ P\subset\mathbf{problem}\}\rangle$$

In order for the definite NP to refer there must be a unique man that has problems. But notice that strengthening **problem** to some P, $P\subset$**problem**, would either pick out the same man, if that man has problem P, or lead to a non-referring description, if he doesn't. Hence no alternative P can ever lead to a stronger proposition. This is different in (91.b):

(93) $$\langle\lambda Q\lambda i\neg\exists x\exists y[\mathbf{go.out.with}_i(\mathbf{m},x) \wedge \mathbf{man}_i(x) \wedge Q_i(y) \wedge \mathbf{have}_i(x,y)],\ \mathbf{problem},\ \{P|\ P\subset\mathbf{problem}\}\rangle$$

Note that in this case choosing stronger alternatives may lead to a stronger overall proposition; for example, the set of worlds in which Mary doesn't kiss men with a specific problem Q may be a subset of the set of worlds in which Mary doesn't kiss men with any problems at all.

7.2. Narrow-Scope Illocutionary Operators

Another contrast that seems to call for a syntactic theory is illustrated with the following pair of examples, illustrating the difference between so-called bridge verbs and non-bridge verbs:

(94) a. Mary didn't think that John had any problems.

b. ??Mary didn't shout that John had any problems.

This contrast can be explained by assuming that non-bridge verbs like *shout* are essentially quotational and hence embed a structure that contains an illocutionary operator. Hence the cases (94.a,b) are analyzed as follows:

(95) a. **Scal.Assert**[Mary didn't think that John had any problems]

b. **Assert**[Mary didn't shout that **Scal.Assert**[John had any problems]]

We can derive that (94.b) is bad as follows: The non-bridge verb *shout*, being quotational, enforces the presence of some illocutionary operator on the embedded sentence. In (94.b), this operator is applied to a BFA structure induced by a polarity item, hence it must be **Scal.Assert**, but the pragmatic requirements for **Scal.Assert** are evidently not satisfied.

The notion of 'quotational' verbs should not be understood in a too narrow sense. For example, the verb *say* certainly can be used in a quotational sense, but also in another sense where only the information content, but not the actual wording is reported by the embedded sentence. Conseqently, *say* is a bridge verb and is transparent for the licensing of polarity items, as in *Mary didn't say that John had any problems*. The analysis of non-bridge verbs as involving embedded illocutionary operators has been proposed for a different set of facts by Song (1994).

Linebarger (1986) has drawn attention to the fact that a NPI must be in the immediate scope of a licensing negation. This explains reading differences like the following:

(96) a. Mary didn't show every child a picture.
  i) Not (Every child x (Some picture y (Mary showed y to x)))
  ii) Not (Some picture y (Every child x (Mary showed y to x)))

b. Mary didn't show every child any picture.
  i) *Not (Every child x (Some picture y (Mary showed y to x)))
  ii) Not (Some picture y (Every child x (Mary showed y to x)))

The absence of reading (i) for (96.b) is unexpected: The quantifier EVERY is upward entailing in its nuclear scope, hence downward entailing under negation. We can explain the lack of reading (i) by assuming that the nuclear scope of a quantifier always is a locus of exploitation for polarity items. Evidence for that comes from the fact that we do find PPIs in cases like the following:

(97) John didn't give every child tons of money.

If the scopal orderings of negation, universal quantifier and any-phrase illustrated in (96.b) are possible (which is shown by 96.a), then we get semantic representations that can be illustrated as follows:

(98) b'. i) **Assert**[Not(every child x (**Scal.Assert**[any picture y (Mary showed x to y))])]

ii) **Scal.Assert**[Not(any picture y (every child x (**Assert**[Mary showed x to y]))))]

Notice that (i) is bad, as the BFA-proposition represented by „any picture y (Mary showed x to y)“ violates the conditions of **Scal.Assert**, just like the sentence *Mary showed John any picture* would do. On the other hand, (ii) is good, as it yields a BFA-proposition that satisfies the felicity conditions of **Scal.Assert**.

8. Interrogatives

One of the most serious problems of most existing accounts of NPIs is that they fail to explain why NPIs occur in questions. We find NPIs in both rhetorical questions and information questions (cf. Borkin 1971):

(99) a. Did Mary ever lift a finger to help you?

b. Who ever lifted a finger to help you?

(100) a. Have you ever been to China?

b. Which student has ever been to China?

Ladusaw (1979) tried to explain the occurrence of NPIs by adopting a principle that a question should be worded in a way that facilitates the answer. Hence if the speaker expects a negative answer, he may use an NPI. This may be an explanation for the occurrence of NPIs in rhetorical questions like (99.a,b), but doesn't apply to information questions like (100.a,b).

A more promising account for NPIs in questions is due to Fauconnier (1980). Fauconnier studies in particular indirect questions embedded under wonder, as I wonder whether this knife can cut even the most tender meat. He observes that, if we restrict our attention to propositions that are considered possible, we have the following implication reversal:

(101) For all ϕ, ψ that are considered (at least) possible:

If $\phi \rightarrow \psi$, then *I wonder whether* ψ entails: *I wonder whether* ϕ.

For example, *I wonder whether John ate a vegetable*, which can be paraphrased as „I am not sure about whether John ate a vegetable“, entails *I wonder whether John ate a carrot*, i.e. „I am not sure about whether John ate a carrot“, but not vice versa — I may be not sure about whether John ate a carrot, but know that John ate a piece of broccoli. This in turn explains, according to Fauconnier, why we find NPIs in questions. For example, (99.a) indicates that the speaker is not even sure whether Mary made a minimal effort to help you, let alone whether Mary made a bigger effort. And as the proposition in question is the minimal one of its scale, Fauconnier assumes that a negative implicature arises that the speaker is sure that Mary didn't make any effort at all to help you.

While I think that Fauconnier's account is promising, it still needs further elaboration. Even then it seems that it covers only those cases where NPIs occur in rhetorical questions, due to the negative implicature just mentioned.[17]

It should be easier to find an answer within the current pragmatic setting. We have seen that in assertions, polarity items induce potential alternative assertions, which must be licensed by pragmatic principles. So we should assume that polarity items may induce potential alternative questions when they occur within a question, and those alternative questions must be licensed by pragmatic principles. This suggests the following format for the interpretation of questions:

(102) $\mathbf{Quest}(\langle B,F,A\rangle)(c) = \mathbf{Quest}(B(F))(c)$,

where for every alternative F', $F' \in A$, speaker has reasons not to base the question on F', i.e. not to propose $\mathbf{Quest}(B(F'))(c)$.

In Krifka (1990, 1992) I have discussed possible reasons that speakers might have not to ask alternative questions:

— In rhetorical questions, the speaker tries to lower the threshold for a positive answer, showing that he is certain that the answer would be negative. For example, in (99.a) the speaker wants to demonstrate how certain he is that Mary didn't help you at all by making the conditions for a positive answer as weak as possible.

— In information questions, the speaker intends to construct the question in such a way that every suggested answer would roughly yield the same amount of information increase. This principle can be illustrated by a game where one player draws a card from a deck of cards and the other has to guess it with as few questions as possible. It would be uneconomical to start with guesses like *Is it the seven of diamonds?*; it is better to start with questions like *Is it a seven?*, or *Is it a diamonds?* A question like (100.a) indicates that the speaker has a reason to prefer the more general question over any alternative, presumably because his information state is such that he expects a betteroverall information gain from an answer to the more general question.

This line of explanation of NPIs in questions captures a generalization alluded to by Borkin (1971), namely that strong NPIs (i.e. idiomatic NPIs with emphatic stress) typically occur in rhetorical questions, whereas weak NPIs tend to occur in information questions. The purpose of an NPI in a rhetorical question is to signal that the speaker tries to make a positive answer as easy as possible, and therefore it is to be expected that he or she selects a question that is based on a proposition that is „extremely“ weak.

Let me flesh out the theory of NPIs in questions I would like to propose. Take some semantic analysis of questions, for instance Groenendijk & Stokhof (1984), where a question is interpreted as a partition on the set of indices and the cells of the partition correspond to the propositions that are full answers to that question. Within a dynamic theory, a question maps an input state to a set of states, and the corresponding answer takes up such a set and picks out one element. For the simplest case of Yes/No questions and corresponding answers, we can assume the following operators (within a static framework):

(103) a. $\mathbf{YN.Quest}(p)(c) = \{c \cap p, c - p\}$

b. $\mathbf{Answ.Assert}(p)(C) = \cup\{c \cap p | c \in C\}$, if $\exists q[q \in C \wedge q \cap p = \emptyset]$)

A Yes/No question based on a proposition p at a common ground c leads to a set C' of two output common grounds, one where p holds and one where p doesn't hold. In turn, a proposition p is a felicitous answer to a set of information states C if it eliminates at least one possibility contemplated by the question, i.e. one element of C. The information conveyed is the union of all states in C that are not eliminated, updated with p. For example, a question like *Is it raining?* at c yields a set C = $\{c \cap \mathbf{raining}, c - \mathbf{raining}\}$. An answer like *It is raining and snowing* then yields an output

$$[c \cap \mathbf{raining}] \cap [\mathbf{raining} \cap \mathbf{snowing}] \cup [c - \mathbf{raining}] \cap [\mathbf{raining} \cap \mathbf{snowing}]$$

$$= [c \cap \mathbf{raining}] \cap [\mathbf{raining} \cap \mathbf{snowing}] \cup \emptyset$$

$$= [c \cap \mathbf{raining} \cap \mathbf{snowing}].$$

An inappropriate answer like *Grass is green* would be infelicitous, as its proposition is compatible with both elements of C.

For questions based on BFA structures we can assume the following rule:

(104) $\textbf{YN.Quest}(\langle B,F,A\rangle)(c) = \{c\cap B(F), c-B(F)\}$,

iff for all $F'\in A$, speaker has reasons not to propose $\{c\cap B(F'), c-B(F')\}$

If the BFA structure was generated by a NPI, we have that in general, $B(F')\subseteq B(F)$, for all F', $F'\in A$. The speaker may have the following reasons in this case:

— In the case of rhetorical questions there are two theoretical options: Either the speaker is so convinced that the answer will be negative that he maximizes the *a priori* possibility for a positive answer. Note that $c\cap B(F)$ will in general be a superset of $c\cap B(F')$. Or the speaker suggests that the common ground c is such that $c\cap B(F')=\emptyset$, which would trivialize the alternative answers. For example, a question like (99.a) may be uttered with respect to a common ground for which the speaker thinks that it is already established that Mary didn't do anything substantial to help you, and hence has to ask the question whether Mary did something minimal to help you as the only remaining one.

— In the case of information questions, the speaker wants to maintain an equilibrium between the informational value of the positive and negative answer; $c\cap B(F)$ and $c-B(F)$ should have roughly the same probability in c. A stronger question based on alternative propositions B(F') would violate this equilibrium. Note that a question like (100.a) would be inappropriate if it is already known that you have been to China, or if the focus of interest is on whether you have been to China in a certain year.

The analysis of direct questions presented here carries over to indirect questions when we assume that the truth conditions of sentences containing indirect questions embody the felicity conditions of the corresponding direct questions. For example, a question like *I wonder whether Mary has ever been to China* will express that the speaker is in an information state where it would make sense for him (i.e. increase his information in an optimal way) to ask the question *Has Mary ever been to China?*.

9. Conclusion

Let us come to a conclusion. In this article I have tried to show that we can arrive at an explanatory theory of the distribution of polarity items within a framework that claims (a) that polarity items introduce alternatives that lead to an informativity relation with respect to the meanings of the polarity items themselves and the common ground at which they are used; and (b) that illocutionary operators make crucial use of this additional information. Polarity items then are just a special case of other constructions that introduce alternatives, like expressions in focus and expressions that are part of a linguistic scale and introduce scalar implicatures.

Of course, vast areas still have to be filled out in this picture to see whether this approach is on the right track. In particular, the range of polarity items and the various construction types and pragmatic constellations that allow for polarity items remains to be investigated in detail.

Notes

Research leading to this paper was carried out while the author was a visiting scholar at the *Arbeitsgruppe Strukturelle Grammatik der Max-Planck-Gesellschaft an der Humboldt Universität* in Berlin and a guest scientist at the *IBM Germany Scientific Centre* in Heidelberg. It was finished during my residency at the *Center for Advanced Study in the Behavioral Sciences*, Stanford. I wish to express my gratitude to these organizations. Parts of the content of this paper was presented at talks at the University of Massachusetts at Amherst in November 1993, and at the conference SALT 4 („Semantics and Linguistic Theory“) at the University of Rochester in May 1994; I am grateful for the comments that I received from the audiences there. In particular, I wish to thank Gene Rohrbaugh and Jack Hoeksema for helpful suggestions concerning the content and the presentation of this paper.

[1] Hence they are treated like *seek* in Montague (1973). See Zimmermann (1993) for an analysis of opaque predicates in terms of arguments of type $\langle e,g \rangle$. At the current point I will discuss matters in an extensional framework for simplicity.

[2] Other uses of forget allow for non-specific readings, e.g. *John forgot a toothbrush*, in the sense of 'John forgot to bring a toothbrush'. Note that we don't have polarity items in this case, cf. **John forgot anything to read*.

[3] To be sure, *until* has a non-NPI variant, as in *John slept until five*. But the NPI variant is clearly distinct from that. Theories that try to derive a sentence like *John didn't* {*wake up / sleet*} *until five* using non-NPI *until* do not predict that it is part of the conventional meaning of these sentences that John woke up or fell asleep at five. See Karttunen (1974) for this observation and Declerck (1995) for discussion.

[4] Of course, it is well attested that regular NPIs may develop into highly grammaticalized morphemes that accompany clausemate negation. This is happening in French with negation patterns like *ne...pas* > *pas*, *ne...personne* > *personne*, where the second element clearly has the flavour of a regular NPI insofar it denotes small entities or unspecific properties, like a step or a person.

[5] J. Hoeksema (pers. comm.) also mentioned Dutch voor de poes (lit. 'for the cat').

[6] However, strong NPIs do not always carry the main stress of a sentence. In particular, contrastive stress overrided stress on strong NPIs, as in *JOHN didn't lift a finger to help me, not MARy*.

[7] This is different from other implementations, where $F \in A$. It will allow a more succinct formulation of certain rules, but nothing of theoretical importance hinges on this decision.

[8] Krifka (1992) defines a type system for structured meanings and also gives rules for what happens if both α and β are structured meanings, and how structures meanings are used by focus-sensitive operators.

[9] This is in response to a criticism made in Kadom & Landman (1993).

[10] It is important to notice that not every expression denoting a decreased standard of precision will create NPIs; for example, *kind of* and *sort of* do not. The reason is that these expressions do not induce alternatives, but simply indicate a more liberal way of applying predicates.

[11] This assumes a model in which liquids have semantic atoms, which is contrary to certain theories that assume a non-atomic part relation for the semantics of mass nouns. I think that speakers can employ semantic models of varying granularity to a given situation, some of which will impose an atomic model for mass nouns, and that the use of the NPI *a drop* implies such an atomic model.

[12] Within a lattice-theoretic setting, we could even identify such amounts as anti-atoms (x is an anti-atom iff there is no y, $y \neq T$,the top element) such that $x \subset y$).

[13] There is, of course, another use where *rather* expresses preferences, as in *I would rather go home*.

[14] This becomes obvious with various presupposition tests. For example, both *Mary is already here* and *It is possible that Mary is already here* entail, and hence presuppose, that Mary arrives at some time.

[15] Notice that this is a different kind of *rather* than the one treated in section 5. Presumably *would rather* is an idiomatic expression, where the foreground expresses a maximal preference, and the alternatives express non-maximal preferences.

[16] There are certain cases of negated NPIs or PPIs in the protasis of conditionals that cannot be explained by local licensing. One example is the sentence *If John doesn't know ANYthing [at ALL] about logic, he will not know Modus Ponens*. Notice that we have a stressed NPI, which means that we should assume that it is exploited by the outermost illocutionary operator. Notice also that it has to be paraphrased by 'Only if John doesn't know anything at all about logic will he not know Modus Ponens'. This shows that another type of assertion is involved that we haven't discussed so far, a type that we find with contrastive assertions and which we can call „exhaustive“. The pragmatic effect of exhaustive assertion of a proposition $\langle B,F,A \rangle$ is that B(F) is claimed to be the only proposition among the alternatives B(F'), $F' \in A$ that is true. Applied to our example this would mean that the Modus Ponens is considered to be the something that people know even if their knowledge of logic is minimal.

[17] A more recent attempt to explain NPIs in questions can be found in Higginbotham (1993).

Reference

Baker, C. L. (1970). Double Negatives. *Linguistic Inquiry* 1, 16-189.

Borkin, A. (1971). Polarity Items in Questions. *Chicago Linguistic Society* 7, 53-62.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

Declerck, R. (1995). The Problem of „not...until“. *Linguistics* 33, 51-98.

Dowty, David. (1994). The Role of Negative Polarity and Concord Marking in Natural Language Reasoning. *Proceedings of Semantics and Linguistic Theory IV*, 114-144, Cornell University.

Edmondson, J.A. (1981). Affecticity and Gradient Scope. *Chicago Linguistic Society* 17, 38-44.

Fauconnier, G. (1975). Polarity and the Scale Principle. *Chicago Linguistic Society* 11, 188-199.

Fauconnier, G. (1978). Implication Reversal in a Natural Language. In F. Guenthner & S. Schmidt (eds.), *Formal Semantics and Pragmatics for Natural Languages*, 289-301. Dordrecht: Reidel.

Fauconnier, G. (1980). Pragmatic Entailment and Questions. In J. R. Searle, F. Kiefer, & M. Bierwisch (eds.), *Speech Act Theory and Pragmatics*, 57-69. Dordrecht: Reidel.

Gazdar, G. (1979). *Pragmatics: Implicature, Presuposition and Logical Form*. New York: Academic Press.

Grice, P. (1975). Logic and Conversation. In *Syntax and Semantics 3: Speech Acts*, 41-58. New York: Academic Press.

Groenendijk, J., & Stokhof, M. (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. Doctoral Dissertation, Department of Philosophy, University of Amsterdam.

Heim, I. (1987). A Note on Negative Polarity and Downward Entailingness. *Proceedings of NELS 14*, 98-107.

Higginbotham, J. (1993). Interrogatives. In K. Hale & S.J. Keyser (eds.), *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger,* Cambridge, MIT Press, 195-228.

Hoeksema, J. (1986). Monotonicity Phenomena in Natural Language. *Linguistic Analysis* 16, 25-40.

Horn, L. (1978). Some Aspects of Negation. In J.H. Greenberg e.al. (ed.), *Universals of Human Language*, Volume 4: Syntax, Stanford University Press, 127-210.

Jacobs, J. (1983). *Fokus und Skalen. Zur Syntax und Semantik von Gradpartikeln im Deutschen*. Tübingen: Niemeyer.

Jacobs, J. (1984). Funktionale Satzperspektive und Illokutionssemantik. *Linguistische Berichte* 91, 25-89.

Jacobs, J. (1991). Focus Ambiguities. *Journal of Semantics* 8, 1-36.

Kadmon, N., & Landman, F. (1993). „Any“. *Linguistics & Philosophy* 16.

Karttunen, L. (1971). Subjunctive Conditionals and Polarity Reversal. *Papers in Linguistics* 4, 279-298.

Karttunen, L. (1974). „Until". *Chicago Linguistic Society* 10.

Klima, E. (1964). *Negation in English. In The Structure of Language.* Englewood Cliffs: Prentice-Hall.

Krifka, M. (1990). Polarity Phenomena and Alternative Semantics. In M. Stokhof & L. Torenvliet (eds.), *Proceedings of the 7th Amsterdam Colloquium 277-302.* Amsterdam: ITLI.

Krifka, M. (1992). Some Remarks on Polarity Items. In D. Zaefferer (ed.), *Semantic Universals and Universal Semantics*, 150-189. Berlin: Foris.

Krifka, M. (1993). Focus and Presupposition in Dynamic Interpretation. *Journal of Semantics* 10, 269-300.

Krifka, M. (1994). The Semantics and Pragmatics of Weak and Strong Polarity Items in Assertions. *Proceedings from Semantics and Linguistic Theory IV,* Cornell University, 195-220.

Ladusaw, W. (1979). *Polarity Sensitivity as Inherent Scope Relations*. Ph.D. Dissertation, University of Texas at Austin.

Ladusaw, B. (1992). Expressing Negation. *Proceedings of Semantics and Lingustic Theory II,* Columbus. University of Ohio.

Landman, F. (1992). *Structures for Semantics.* Dordrecht: Kluwer.

Levinson, S. C. (1984). *Pragmatics.* Cambridge University Press. Cambridge.

Lewis, D. (1970). General Semantics. *Synthese* 22, 18-67.

Linebarger, M. (1991). Negative polarity as linguistic evidence. *Chicago Linguistic Society 27-II: Papers from the parasession on negation*.

Linebarger, M. C. (1980). *The Grammar of Negative Polarity.* Ph.D. Dissertation, MIT. Cambridge, MA.

Linebarger, M. C. (1987). Negative Polarity and Grammatical Representation Linebarger. *Linguistics and Philosophy* 10, 325-387.

May, R. 1985. *Logical Form. Its Structure and Derivation.* MIT Press, Cambridge, Massachusetts.

Partee, B. H. (1984). Nominal and Temporal Anaphora. *Linguistics and Philosophy* 7, 243-286.

Progovac, L. (1988). *A Binding-Theoretic Approach to Polarity Sensitivity.* Ph.D. Dissertation, University of Southern California, Los Angeles.

Progovac, L. (1992). Negative Polarity: A semantico-syntactic approach. *Lingua* 86, 271-299.

Progovac, L. (1993). Negative Polarity: Entailment and Binding. *Linguistics and Philosophy* 16, 149-180.

Progovac, L. (1994). *Negative and Positive Polarity. A Binding Approach.* Cambridge, Cambridge University Press.

Rohrbaugh, Gene. (1993). Focus and the Licensing of Negative Polarity Items. Ms., University of Texas at Austin.

Rooth, M. E. (1985). *Association with Focus.* Ph.D.-Dissertation, University of Massachusetts at Amherst.

Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics* 1, 75-116.

Song, Jae-Gyun. *Clause-Embedding Verbs and the Interpretation of WH-IN-SITU*. Ph.D. dissetation, University of Texas at Austin.

Stalnaker, R. (1975). Indicative Conditionals. *Philosophia* 5, 269-286.

Stalnaker, R. C. (1972). Pragmatics. In D. Davidson & G. Harman (eds.), *Semantics of Natural Language*, Dordrecht: North-Holland, 380-397.

Stechow, A. v. (1990). Focusing and Backgrounding Operators. In W. Abraham (ed.), *Discourse Particles*, Pragmatics & Beyond. Amsterdam: John Benjamins, 37-84.

Zimmermann, E. (1993). On the Proper Treatment of Opacity in Certain Verbs. *Natural Language Semantics* 1 (1992-1993), 149-180.

Zwarts, F. (1993). Three Types of Polarity. Ms, University of Groningen.