

Negative Polarity and General Statements

Eric Jackson

Corona Corporation

1 Introduction

In this paper, I will present an account of two types of negative polarity items (the so-called “strong” and “weak” types). The structure of the paper is as follows. In section two, I briefly outline the data, the leading accounts, and the main motivation for my account. Section three contains the account itself and application of it to various data. In section four, I mention some formal properties of the proposal and discuss its explanatory power.

2 The Problem

2.1 The Data

The classic example of a weak negative polarity item (NPI) is the determiner *any*. As has often been noted, *any* in many cases seems to prefer negative environments, in some sense, over positive ones:

- (1) I don't have any apples.
- (2) * I have any apples.
- (3) No student has any apples.
- (4) * Every student has any apples.

However, the following data show that the class of NPI licensors include more than just different sorts of negation, at least in any pre-theoretic sense of “negation”:

- (5) * Some student has any apples.
- (6) At most two students have any apples.
- (7) * Most students have any apples.
- (8) Few students have any apples.
- (9) No student who has any apples is happy.
- (10) Every student who has any apples is happy.
- (11) * Some student who has any apples is happy.
- (12) At most two students who have any apples are happy.
- (13) Most students who have any apples are happy.
- (14) Few students who have any apples are happy.

Only a subset of the contexts that license NPIs are illustrated above. Other such contexts include comparatives, conditionals, adversative predicates and questions. However, I will limit my attention here to NPI licensing by quantified noun phrases.

The reader is invited to consult (Jackson 1994) for a fuller discussion of other sorts of NPI licensing environments.

One assumption I am making about *any* that will be important below is that *any* is an existential quantifier, rather than a universal quantifier. In some sentences, such as (1), it is not obvious which of these two analyses is correct. However, in others, such as (6), it is clear that *any* must be functioning as an existential quantifier. (There are cases where *any* functions as a universal quantifier, the so-called “free-choice” uses of *any*, but these can be distinguished from the “polarity-sensitive” use, and will not be considered here.) Other arguments for the existential status of *any* can be found in the literature (e.g., (Ladusaw 1980), (Linebarger 1980)).

As their name suggests, the strong NPIs appear in a more restricted set of contexts than the weak NPIs. Here are examples of the distribution of the Dutch expression *ook maar*, the classic example of a strong NPI:

- (15) Geen kind heeft ook maar iets gezien.
No child has anything seen
No child saw anything.
- (16) Ieder kind dat ook maar iets heeft gezien mag nu naar huis gaan.
Every child that anything has seen may now to house go
Every child that has seen anything may now go home.
- (17) * Hoogstens twee kinderen hebben ook maar iets gezien.
At most two children have anything seen
At most two children saw anything.
- (18) * Weinig kinderen hebben ook maar iets gezien.
Few children have anything seen
Few children saw anything.
- (19) * De meeste kinderen die ook maar iets hebben gezien mogen nu naar huis gaan.
Most children that anything have seen may now to house go
Most children that saw anything may now go home.

2.2 Leading Accounts

The leading account of the weak NPIs is that of Ladusaw (Ladusaw 1980). The central idea is very simple: an NPI must be licensed by a monotone decreasing operator. A monotone decreasing operator is one such that:

$$O(A) \wedge B \subseteq A \Rightarrow O(B)$$

Thus, for example, (20) is licensed because it entails (21):

- (20) No student has any apples.

(21) No student has any red apples.

The reader can verify that this account makes the right prediction for each of the sentences above, other than (13) and (14). To see that (13) is problematic, note that (22) does not entail (23):

(22) Most students that ate any apples are happy.

(23) Most students that ate any poisoned apples are happy.

The leading account of the strong NPIs is that of Frans Zwarts (Zwarts 1986, Zwarts 1993). It is like Ladusaw's account, in that the items in question are claimed to be licensed by operators with a certain logical property. In this case, the relevant class of operators are the anti-additive ones. An anti-additive operator is one such that:

$$O(A) \Leftrightarrow \bigwedge_i O(A_i) \quad (A = \cup_i A_i)$$

As with downward monotonicity, we can check for anti-additivity by seeing if certain entailments are valid. For example, the anti-additivity of *nobody* can be determined by checking that the following hold:

- Nobody (VP_1 or VP_2) \Rightarrow Nobody VP_1 and nobody VP_2
- Nobody VP_1 and nobody $VP_2 \Rightarrow$ Nobody (VP_1 or VP_2)
- Nobody VP_1 or nobody $VP_2 \Rightarrow$ Nobody (VP_1 and VP_2)

The Zwarts analysis also makes the correct predictions for the other environments considered above.

The reader is referred to the work of Ladusaw and Zwarts for a fuller discussion of their accounts.

2.3 Motivation

2.3.1 Strength and Generality

The guiding intuition for my account will be an intuition that many have had in one form or another; namely, that negative polarity items are related in some way to what we might call "strong statements". This intuition is most explicitly incorporated in the account of Kadmon and Landman (Kadmon and Landman 1993), who posit that NPIs induce strengthening of a statement. This intuition can be supported by comparing pairs of sentences like the following:

(24) a. John didn't lift a finger.

b. John did nothing.

(25) a. John didn't donate a red cent.

b. John donated no money.

(26) a. John didn't budge an inch.

b. John didn't move.

Although (24a) and (24b) are roughly synonymous, at least in terms of truth conditions, it seems that (24a) has a special emphatic flavor that is lacking in (24b). Likewise for (25a) and (25b) and (26a) and (26b). This emphatic flavor can be found in many other NPIs besides the idioms. Taking *any*, for example, consider the following sentences, which were discussed by Kadmon and Landman:

- (27) a. I don't have any potatoes.
 b. I don't have potatoes.
 c. I don't have a potato.

(27a) seems to be stronger or more emphatic, in some sense, than (27b) or (27c). Along these lines, note what happens if we focus the occurrences of *any* and *a* in (27a) and (27c):

- (28) I don't have ANY potatoes.
 (29) I don't have A potato.

With *any*, focusing increases the emphatic flavor, whereas with *a* focusing has no such effect. The question that requires explanation, then, is why do so many NPIs have this emphatic flavor, this association with strong statements. It is not apparent what downward monotonicity, for example, has to do with strength. In contrast, an account such as Kadmon and Landman's seems to have more potential for explaining this observation.

However, I believe Kadmon and Landman's account goes astray in the way they try to cash out the intuition just discussed. They claim that *any* introduces *truth-conditional* strengthening over corresponding sentences with other indefinites such as NPs with *a*, or bare plurals. This position is perhaps tenable for examples like (27), but it plainly is not for numerous other examples, such as:

- (30) a. John hasn't visited any African countries.
 b. John hasn't visited an African country.

(30a) and (30b) are clearly equivalent; hence there is no truth-conditional strengthening induced by *any*.

The moral I draw from examples like this is that we should not attempt to define "strong statement" in a "relative" way, as Kadmon and Landman do, by insisting that statements with NPIs have different truth conditions than corresponding statements without NPIs. Instead, I will propose an "absolute" definition. In this way I will not prohibit statements without NPIs from being strong, and, hence, the equivalence of (30a) and (30b) will not be a problem for me.

It is worth taking a moment here to reexamine whether the notion we need to capture is precisely that of a *strong* statement. Presumably, strength is related to entailment relations: one statement is stronger than another if the first entails the second. With this in mind, consider the following:

- (31) No student has any apples.
 (32) * Exactly two students have any apples.

(33) Every student who has any apples is happy.

(34) * Exactly two students who have any apples are happy.

In these minimal pairs, it is not the case that the NPI-licensing sentence entails the sentence that does not license NPIs. Hence, it is not clear how one would go about defining “strong statement” in such a way that we get the right result for sentences like (31–34). Thus, I will instead propose that NPIs are licensed in what I will call “general statements”. At an intuitive level, (31) and (33) are “general” because they make universal, or negative universal, claims about students and apples. A sentence like (6) may not be strictly general, but it does seem “close” to general, in the sense that it entails that there is *almost* no apple ownership among students.

2.3.2 Intervention Effects

One empirical problem for many theories of negative polarity involves what we might term “intervention” effects (see (Linebarger 1980, Linebarger 1987)). For theories that posit that NPIs are licensed by a certain sort of operator (like the Ladusaw and Zwarts theories), the problem is to account for why certain sorts of things seem to block NPI licensing. For example, contrast the following ((35) is from (Linebarger 1987)):

(35) * John didn’t give a red cent to every charity.

(36) John didn’t give a red cent to any charity.

A similar contrast can be found among the following:

(37) * No student gave every teacher any apples.

(38) No student gave a teacher any apples.

(39) No student gave more than a few teachers any apples.

The ‘*’ for (37) indicates that it is unacceptable under the (pragmatically favored) reading in which *every teacher* outscopes *any apples*. If (37) seems less bad than (35), I suspect it is simply because it is hard to prevent oneself from giving *any apples* wide scope over *every teacher*.

These data are problematic for the original Ladusaw analysis because the NPI occurs within the scope of a monotone decreasing operator, and yet the sentence is unacceptable. Some work has suggested that any quantified expression will block NPI licensing, but (38) and (39) are counterexamples to that claim. It should also be noted that couching Ladusaw’s theory in terms of (possibly complex) *environments*, instead of operators, will not help. *No student gave every teacher X* is a monotone decreasing environment, as evidenced by the entailment of (41) by (40):

(40) No student gave every teacher an apple.

(41) No student gave every teacher a red apple.

Thus, the unacceptability of (35) and (37) is problematic for this hypothetical account.

It is interesting to note that the contrast between (37) and (38) is paralleled in a contrast between the modals *might* and *must*:

(42) Everyone who might have ever met the victim was investigated.

(43) * Everyone who must have ever met the victim was investigated.

Might is like an existential quantifier in that you can think of it as an existential quantifier over possible worlds, while *must* is like a universal quantifier in that you can think of it as a universal quantifier over possible worlds.

The moral I draw from examples like (35–39) and (42–43) is that the class of operators that cause intervention effects is determined by a semantic principle or principles — just like the class of operators that license NPIs. The question, though, is how to achieve this in a natural way. In particular, it seems unlikely that there are two unrelated semantic accounts, one for the licensing operators, and one for the intervening operators. In the account I present, the class of licensing and intervening operators will fall out from a single semantic principle.

2.3.3 Unanswered Questions

Further motivation for my account comes from puzzling questions that I believe cannot be answered under the Ladusaw/Zwarts theory. The first is “Why do so many NPIs correspond to existential quantifiers?” We have argued already that polarity-sensitive *any* is an existential quantifier. For similar reasons, it should also seem clear that *ever* is an existential quantifier (over time). Consider also several examples of the negative-polarity idioms:

(44) Mary didn’t lift a finger.

(45) Mary didn’t do a thing.

(46) Mary didn’t hear a sound.

(47) This painting isn’t worth beans.

(48) John didn’t do (jack)shit.

All of these idioms involve an item that is typically translated as an existential quantifier. In (44–46), it is a singular indefinite, in (47) it is a bare plural, and in (48) it is a mass noun. Hence it is not surprising that it turns out to be plausible to interpret these expressions — in their idiomatic interpretations, as well as their literal interpretations — as existentially quantified expressions. The idiomatic reading of (44), for example, can be given the following logical form:

(49) \neg some x (action(x), did(Mary, x))

Even some NPIs that may not seem at first sight to be existential quantifiers turn out to be naturally analyzed as such upon closer inspection. Consider, for example, *yet*:

(50) No one is here yet.

(51) * Everyone is here yet.

At first glance, one might think that *yet* simply means *now*, so that (50) means that no one is now here. However, note that (50) would be an extremely odd thing to say at a party, say, if someone had arrived a few minutes ago and then departed just before the utterance of the sentence. (50) seems to, in fact, convey the idea that no one has been here during a contextually determined interval of time ending at the present moment. This reading results if we treat *yet* as an existential quantifier over points of time in this interval. So, for example, the logical form of (50) would be (52) (where t_0 is the contextually determined start point of the interval):

(52) $\text{no } x(\text{person}(x), \text{some } t(t < \text{now} \wedge t > t_0, \text{is-here-at}(x, t)))$

Consider also *until*. As Karttunen (Karttunen 1974) has argued, this item has a polarity-sensitive use (as well as a non-polarity-sensitive use). Only the polarity-sensitive use is available when an *until* clause combines with a non-durative predicate. For example:

- (53) * John left until the host spilled beer on him.
- (54) John didn't leave until the host spilled beer on him.
- (55) * Everyone left until the host spilled beer on him.
- (56) No one left until the host spilled beer on him.
- (57) Few people left until the host spilled beer on them.

As Karttunen points out, on its polarity-sensitive usage *until* means the same as *before*. This can be checked by replacing *until* with *before* in (54), (56) and (57). Furthermore, *before* (and hence this meaning of *until*) can be naturally treated as an existential quantifier over time. On this analysis, (56) gets the following logical form:

(58) $\text{no } x(\text{person}(x), \text{some } t(\text{every } t'(\text{spilled-beer-on-at}(\text{the-host}, x, t'), t < t'), \text{left-at}(x, t)))$

The relevance of the generalization I am arguing for in this section is the following. If polarity-sensitivity is just a lexical property that items are free to have or not to have, then there is no apparent reason why universal quantifiers should never (or almost never) be NPIs. In general, the Ladusaw and Zwarts accounts give no reason to believe that polarity sensitivity should be associated with some types of items more than others.

A related puzzle for any theory of polarity is why there are the limited range of NPI licensing conditions that there are. We have seen two types, the strong and the weak, and there may be others, but there does not seem to be an unlimited number. For the Ladusaw/Zwarts approach, the puzzle can be posed via the following questions. Why should there be items that appear only in monotone decreasing contexts, but not items that appear only in monotone increasing contexts? What do downward monotonicity and anti-additivity have in common? In general, we want to know what the various NPI licensing conditions have in common that is not possessed by other possible licensing conditions.

3 The Account

3.1 The Basics

The account I will propose can be seen as an attempt to formalize the idea that NPis occur in general statements. The idea I will pursue is that a statement is general if it is easy to find counterexamples to it (or “negative witnesses”, as I shall term them). With this in mind, consider the following two sentences:

- (59) No student has any apples.
 (60) Every student who has any apples is happy.

Intuitively, (59) is falsified by a tuple of a student and an apple in the ownership relation, and (60) is falsified by a tuple of an unhappy student and an apple in the ownership relation. If we look at other examples, though, we see that our notion of “negative witness” must be defined in such a way as to include small *sets* of tuples that falsify a sentence:

- (61) At most two students have any apples.

A set of three student-apple pairs would be a negative witness for (61), whereas singleton negative witness sets suffice for (59) and (60). Contrast (62):

- (62) * Every student has any apples.

A small set of tuples in the ownership relation cannot falsify (62).

The condition on NPI licensing that I propose is called the “negative witnessing condition”. It holds of an arbitrarily complex formula with respect to an occurrence of an atomic formula within that complex formula. For example, assuming the logical form of (59) is (63), the negative witnessing condition will hold of (63) with respect to the occurrence of *has*:

- (63) no $x(\text{student}(x), \text{some } y(\text{apple}(y), \text{has}(x, y)))$

The condition is as follows:

ϕ satisfies negative witnessing with respect to P iff:

$$\forall M \exists W (\forall w \in W (Small(w)) \wedge \forall A (M(P | A) \models \neg \phi \leftrightarrow \exists w \in W (w \subseteq A)))$$

$M(P | A)$ refers to the model that differs from M at most in that the denotation of P is set to A . What this condition says is that for any model M there is a set of (potential) negative witnesses W . (Ignore, for the moment, the question of what $Small(w)$ is supposed to mean.) As you vary the denotation of the predicate P (holding everything else fixed), the formula ϕ is false if and only if some negative witness is a subset of the denotation of P . For example, for (63) an appropriate choice for the set W (in any given model) is:

$$\{\{ \langle s, a \rangle \} \mid \text{student}(s) \wedge \text{apple}(a)\}$$

We still need to explicitly state the condition that relates NPI licensing to negative witnessing. Here is one way to do this:

An occurrence of an NPI within a sentence is acceptable iff the matrix of the quantifier introduced by the NPI is an atomic formula $P(\dots)$ and some subformula of the logical form of the sentence satisfies negative witnessing with respect to P .

Obviously this will have to be revised somewhat, if only to take care of the case where the matrix of the NPI is not atomic, but it will serve as a starting point. (See (Jackson 1994) for an improved version.)

Now contrast (59) with (62), whose logical form, given below, we should expect to not satisfy the negative witnessing condition:

(64) every $x(\text{student}(x), \text{some } y(\text{apple}(y), \text{has}(x, y)))$

Suppose the negative witnessing condition were satisfied for (64). Consider the following two models M and M' :

$$\begin{array}{ll} \text{student}^M = \{s1, s2\} & \text{student}^{M'} = \{s1, s2\} \\ \text{apple}^M = \{a1\} & \text{apple}^{M'} = \{a1\} \\ \text{has}^M = \{\langle s1, a1 \rangle\} & \text{has}^{M'} = \{\langle s1, a1 \rangle, \langle s2, a1 \rangle\} \end{array}$$

Since (64) is false in M , $\{\langle s1, a1 \rangle\}$ or \emptyset would have to be a member of the set of potential negative witnesses, W , for M . But M' differs from M only on has and $\{\langle s1, a1 \rangle\}$ and \emptyset are both subsets of $\text{has}^{M'}$, so, by the negative witnessing condition, (64) must be false in M' . But (64) is, in fact, true in M' , so the negative witnessing condition is not satisfied.

By a similar argument, one can see that *any* is predicted to be licensed in simple negative sentences (e.g., (1)), but not in simple positive sentences (e.g., (2)).

3.2 Strong NPIs

Before proceeding, we should recall the more restricted distribution of the strong NPIs and our desire to relate the two types of NPI. A natural way to do this would be to have a parametrized negative witnessing condition; under one parameter setting, the condition would predict the distribution of the strong NPIs, while another setting would yield the broader distribution of the weak NPIs. This is, in fact, possible within the approach I am presently describing. In particular, the parametric aspect will be the interpretation of *Small* in the negative witnessing condition. For the strong NPIs, the interpretation of *Small*(w) that I propose is very simple:

For strong NPIs, *Small*(w) means $|w| \leq 1$.

Let us now determine some of the predictions made for strong NPIs. The explanations we provided for the licensing (or lack of licensing) of *any* in (59) and (62) carry over essentially without modification to their Dutch counterparts with *ook maar*, so I will not examine them again. Let us consider instead:

- (65) Ieder kind dat ook maar iets heeft gezien mag nu naar huis gaan.
 Every child that anything has seen may now to house go
 Every child that has seen anything may now go home.

(66) every $x(\text{child}(x) \wedge \text{some } y(\text{thing}(y), \text{has-seen}(x, y)), \text{may-now-go-home}(x))$

The question is whether (66) satisfies negative witnessing with respect to *has-seen*. It does, because the following is a suitable choice for W , as the reader can verify:

$\{\{ \langle a, b \rangle \} \mid \text{child}(a) \wedge \neg \text{may-now-go-home}(a) \wedge \text{thing}(b) \wedge \text{has-seen}(a, b)\}$

I leave it to the reader to verify that the account predicts correctly the acceptability or unacceptability of the other sentences with strong NPIs that we looked at in section 2.1. Let us consider instead some of the intervention effects, such as the contrast between logical forms of the following two types:

(67) $\text{no}(\phi, \text{some}(\chi, \text{some}(\psi, P)))$

(68) $\text{no}(\phi, \text{every}(\chi, \text{some}(\psi, P)))$

First, here is an example corresponding to (67):

- (69) Niemand geeft een kind ook maar iets waaraan het zich kan bezeren.
 No one gives a child anything with which he himself can hurt
 No one gives a child anything with which he can hurt himself.

(70) $\text{no } x(\text{person}(x), \text{some } y(\text{child}(y), \text{some } z(\text{thing}(z) \wedge \text{can-hurt-with}(y, y, z), \text{gives}(x, y, z))))$

It is easy to see that the following is a suitable choice for W for (70):

$\{\{ \langle a, b, c \rangle \} \mid \text{person}(a) \wedge \text{child}(b) \wedge \text{thing}(c) \wedge \text{can-hurt-with}(b, b, c) \wedge \text{gives}(a, b, c)\}$

Contrast the situation with (71) (which corresponds to (68)):

- (71) * Niemand geeft ieder kind ook maar iets waaraan het zich kan bezeren.
 No one gives every child anything with which he himself can hurt
 No one gives every child anything with which he can hurt himself.

(72) $\text{no } x(\text{person}(x), \text{every } y(\text{child}(y), \text{some } z(\text{thing}(z) \wedge \text{can-hurt-with}(y, y, z), \text{gives}(x, y, z))))$

To see that (72) does not satisfy negative witnessing, consider the following model:

$\text{person}^M = \{p1\}$

$\text{child}^M = \{c1, c2\}$

$\text{thing}^M = \{t1\}$

$\text{can-hurt-with}^M = \{\langle c1, c1, t1 \rangle, \langle c2, c2, t1 \rangle\}$

$\text{gives}^M = \{\langle p1, c1, t1 \rangle, \langle p1, c2, t1 \rangle\}$

Suppose negative witnessing were satisfied. Since (72) is false in M , some subset S of $\{\langle p1, c1, t1 \rangle, \langle p1, c2, t1 \rangle\}$ such that $|S| \leq 1$ would have to be in the set of negative witnesses W . But that would mean that the variant of M , M' , where $gives^{M'} = S$ would have to make (72) false too. But, in fact, whether $gives^{M'}$ is $\{\langle p1, c1, t1 \rangle\}$, $\{\langle p1, c2, t1 \rangle\}$ or \emptyset , (72) is true in M' . So negative witnessing is not satisfied.

3.3 Weak NPIs

Now let's return to the weak NPIs, and try to see how *Small(w)* must be interpreted. It is clear that the interpretation must be relaxed, in order to allow for the acceptability of sentences such as the following:

(73) At most two students have any apples.

(74) Few students have any apples.

However, simply increasing the numerical limit on the size of negative witness sets clearly will not suffice. Consider (74). As the number of students grows, so does the size of a negative witness set for (74). If the meaning of (the proportional sense of) *few* is *less than n%*, then the size of negative witness sets for (74) is $n\%$ of the number of students. Therefore, rather than interpreting *Small* as *less than n*, for some fixed n , it would seem preferable to interpret it as follows:

For weak NPIs, *Small(w)* means *less than 50% of the size of* $\{x \mid \phi(x)\}$
for some subformula, $\phi(x)$, of the logical form of the sentence.

I choose 50% as the relevant figure somewhat arbitrarily. However, it seems the figure must be at least this high to allow for the licensing of NPIs by *most* (see below). Probably, though, it would be better to give a graded definition, by which I mean one that would predict diminishing degrees of NPI acceptability as the size of the negative witness sets gets larger. (Some confirmation for this comes from the fact that there is some disagreement in the literature over whether NPIs can be licensed by *most*.)

Turning our attention to the determiner *at most n*, we can find some independent support for the claims that a) NPI licensing is related to smallness of negative witness sets (rather than, say, downward monotonicity), and b) that smallness is to be defined proportionally:

(75) At most one hundred Americans have any children.

(76) ? At most one hundred people in this room have any children.

(76) seems to be less acceptable than (75). This contrast supports claim (a) above, because *at most n* is always monotone decreasing, no matter what *emphn* is. It supports claim (b), because n is the same numerically in each case (100), but not in a proportional sense; i.e., 100 is a relatively small percentage of the total number of Americans, but not of the number of people in a typical room.

Although I do believe that there is a contrast between (75) and (76), it also seems clear that (76) is not completely out, which is a *prima facie* problem for my account. I think that the reason for this is that *at most n* is conventionally used with small *n*, and for the purpose of expressing negative generalizations, and so even when *n* is large, *at most n* has the same NPI licensing properties as it has when *n* is small. In short, I relate the NPI licensing capabilities of *at most n* to an implicature associated with it (that *n* is small). At first glance, this seems like a somewhat *ad hoc* move to make. It is preferable to avoid appeals to pragmatics to save one's theory from difficulties. However, a brief digression to discuss the determiner *exactly n* will show that this appeal to pragmatics is necessary on independent grounds.

The standard assumption in the literature is that *exactly n* does not license NPIs. And, indeed, sentences like (77) seem pretty bad:

(77) * Exactly two students have any apples.

This fact is predicted both by Ladusaw's account (*exactly n* is not monotone decreasing) and by mine. To see that (77) does not satisfy the negative witnessing condition, note that the sentence is false when the ownership relation is empty. If the negative witnessing condition were satisfied, then the empty set would have to be a negative witness set. But, then, since the empty set is a subset of every set, it would follow that (77) must be vacuously false. This is clearly not the case, so (77) does not satisfy the negative witnessing condition.

However, not every sentence with *exactly n* is as bad as (77). Consider the following example, drawn from (Linebarger 1987):

(78) ? Exactly four people in the world have ever read that dissertation: Bill, Mary, Tom and Ed.

Linebarger claims this sentence is acceptable. I have glossed it with a '?', but I agree that the sentence is not too bad, which is a *prima facie* problem for my account and for Ladusaw's. The point I wish to focus on, however, is that judgments markedly improve if *exactly n* is replaced with *just n*:

(79) Just four people have ever read that dissertation: Bill, Mary, Tom and Ed.

Also contrast (77) with (80):

(80) Just two students have any apples.

It seems clear that *just n* has the same truth conditions as *exactly n* in these sentences (and not the same as, say, *at most n*). The relevant difference between the two expressions apparently must be the fact that *just n* implicates that *n* is small, whereas *exactly n* does not. Note furthermore that the *n* in *just n* need not be small in any absolute sense, or even relative to the size of the set quantified over, for NPIs to be licensed:

(81) Just two hundred million Americans have any children.

Although two hundred million is both large in an absolute sense, and large relative to the total number of Americans, (81) is acceptable. Apparently, the implicature contributed by *just n* that *n* is small overrides background expectations.

In general, the contrast between *exactly n* and *just n* seems to force us to the conclusion that a) NPI licensing is sensitive to implicatures, b) one relevant type of implicature is an implicature that some set is small, and c) in such a case, the relevant set need not be small relative to background expectations. What is significant about these conclusions is that they provide independent support for the account of *at most n* that I put forward above. There I suggested that sentences with *at most n* might license NPIs better than my theory strictly predicts because *at most n* is conventionally associated with the use of small *n* for the purpose of expressing negative generalizations. The contrast between *just n* and *exactly n* shows that we need to appeal to precisely this mechanism to explain the behavior of *just n*; hence it is not *ad hoc* to appeal to it in accounting for the facts involving *at most n*.

I claim, then, that the theory I am developing can account at this point for sentences such as the following:

(82) At most two students have any apples.

(83) At most two students who have any apples are happy.

(84) Few students have any apples.

However, more work is needed to handle superficially similar sentences such as (85) and (86):

(85) Few students who have any apples are happy.

(86) Most students who have any apples are happy.

At first glance, it might seem that a set of *n* student-apple pairs could serve as a negative witness for (85) (where *n* is a small percentage of the number of students who have any apples). But, in fact, as the ownership relation varies across different models $M(P \mid A)$, so may the number of apple-owning students. Thus, the set of *n* student-apple pairs that serves as a negative witness in one model may not be big enough to serve as a negative witness in another model. In general, for any model *M*, there is not going to be a single set, *W*, of negative witnesses that will give the right results.

Recall that the idea of $M(P \mid A)$ is that it denotes a model that differs minimally from *M* (viz., it differs only on *P*). The above remarks might be taken as showing that, contrary to intent, the interpretation of *few* varies between *M* and $M(P \mid A)$. Suppose our condition were reformulated so that rather than considering $M(P \mid A)$ for any *A*, we only consider variants of *M* that have the same number of apple-owning students. (The ownership relation would still be able to vary — just not as much as before.) This would result in a fixed interpretation of *few* — *few* would become like *at most n* in all essential respects. Thus, not surprisingly, there would now be a set of negative witnesses *W* for each model, and the negative witnessing condition would be satisfied.

For purposes of comparison, it is worth noting that (85) and (86) are also problematic for Ladusaw's theory. *Most* and *few* are not downward monotonic

within their restrictors.

3.4 More on Intervention Effects

To conclude this section, let me point out an interesting contrast in the intervention effects associated with weak and strong NPIs. For a weak NPI like *any*, we predict (correctly) that a sentence like the following is acceptable (with the exhibited quantifier scoping):

(87) No teacher told more than a few students anything they could find useful.

(88) $\text{no } x(\text{teacher}(x), \text{more-than-a-few } y(\text{student}(y), \text{some } z(\text{can-find-useful}(y, z), \text{told}(x, y, z))))$

(See (39) in section 2.3.2 also.) To see this, note that a set of a few tuples of a teacher, a student and a useful piece of information in the *tell* relation will serve as a negative witness set. Now note that the corresponding sentence in Dutch with *ook maar iets* is not acceptable:

(89) * *Geen docent vertelde meer dan een paar studenten ook maar iets waar ze wat aan hadden*

No teacher told more than a few students anything that they could find useful

No teacher told more than a few students anything they could find useful.

This is what we expect since the negative witness sets are not small enough (i.e., not ≤ 1) to fulfil the condition that licenses strong NPIs.

These data are especially interesting because they support the view that NPIs are licensed by appearing in a (possibly complex) environment with a certain semantic property, rather than by being in the scope of an operator with a certain semantic property (as in the Ladusaw analysis). There have been attempts to capture intervention effects within this latter approach, but these attempts have not, in general, distinguished among different sorts of interveners on semantic grounds, which seems necessary. One could, of course, offer a semantic definition of possible interveners, but it is surely preferable to have a single semantic condition on licensing environments, rather than a condition on possible licensors and a separate condition on possible interveners. In addition to reasons of parsimony, there is the additional consideration that on a “single condition” approach, it is predicted that the strong NPIs should have a broader class of interveners than the weak NPIs. On a “two condition” approach, it would be possible for one type of NPI to have both a narrower class of licensors and a narrower class of interveners than another type.

4 Discussion

4.1 Relationship of Negative Witnessing to Anti-Additivity and Downward Monotonicity

A natural question to ask is what the relationship is between negative witnessing, on the one hand, and anti-additivity and downward monotonicity, on the other. It turns out that there is a simple and illuminating answer. Recall that negative witnessing is a condition on arbitrarily complex formulas, and not merely on operators. If we like, we may think of it as a condition on the environment of an atomic formula. Therefore, we will be interested in anti-additivity and downward monotonicity as conditions on environments as well. We can express these conditions as follows.

ϕ is monotone decreasing with respect to P iff:

$$\forall M, A, B (M(P \mid A) \models \phi \wedge B \subseteq A \rightarrow M(P \mid B) \models \phi)$$

ϕ is anti-additive with respect to P iff:

$$\forall M, A, C (A = \cup C \wedge C \neq \emptyset \rightarrow (M(P \mid A) \models \phi \leftrightarrow \forall c \in C (M(P \mid c) \models \phi)))$$

For easy reference, here is the definition of negative witnessing again. ϕ satisfies negative witnessing with respect to P iff:

$$\forall M \exists W (\forall w \in W (Small(w)) \wedge \forall A (M(P \mid A) \models \neg \phi \leftrightarrow \exists w \in W (w \subseteq A)))$$

As discussed earlier, this is a parameterized condition. The results below involve the following two versions of this condition:

Strict negative witnessing: $Small(w)$ means $|w| \leq 1$.

Unrestricted negative witnessing: $Small(w)$ places no restriction on w .

These are, in some sense, the strongest and weakest reasonable versions of the negative witnessing condition. Note that strict negative witnessing is precisely the condition we associated with strong NPI licensing, but that unrestricted negative witnessing is weaker than the condition we associated with weak NPI licensing.

The relationship to downward monotonicity and anti-additivity is the following:

ϕ is monotone decreasing with respect to P iff ϕ satisfies unrestricted negative witnessing with respect to P.

ϕ is anti-additive with respect to P iff ϕ satisfies strict negative witnessing with respect to P.

See (Jackson 1994) or (Jackson 1995) for the proofs.

Perhaps the most interesting aspect of these results is that they show that we can view downward monotonicity and anti-additivity as two endpoints on a

scale. They correspond to the strongest and the weakest reasonable versions of the negative witnessing condition. As we vary the interpretation of $Small(w)$, we can also produce various conditions with intermediate strength.

A convenient by-product of these results is that we can take advantage of results established in the literature in the Ladusaw/Zwarts tradition. For example, it has been shown that comparatives (Hoeksema 1983) and *before* clauses (Valencia et al. 1994) are anti-additive environments. From the equivalence of anti-additivity and strict negative witnessing, we know that our theory also predicts the licensing of weak and strong NPIs in these contexts.

4.2 Explanatory Power

At this point I would like to return to the problems relating to explanatory power that were raised in the beginning of this paper. The first problem was to explain why negative polarity items seem to take part in especially strong, or general, statements. Along those lines, I would claim that the negative witnessing condition is a natural characterization of the intuitive idea of a general statement. Essentially a negative witness is a set of tuples that serves as a counterexample to a statement. It seems natural to equate the notion of a general statement with the notion of a statement that is “easily counterexemplified”. The second question was why there are the types of NPIs that there are. The assumption that will allow us to answer this question is that negative polarity should be related to general statements, and, hence, negative witnessing. For this to be a reasonable assumption, it must be the case that general statements form a natural class, and, moreover, one which it is reasonable to expect to be marked via NPIs. I believe this should be fairly intuitive. The basic idea is just that from a functional perspective, it is natural to suppose that general statements should be especially useful and especially common. And, in turn, it is not surprising that a type of statement that is especially useful or common should be flagged by means of certain items. Given this, we can see why the great majority of negative polarity items seem to be existential quantifiers. Simply put, most quantifiers other than the existential quantifier cannot take part in formulas that satisfy negative witnessing. Two exceptions are the universal quantifier and the negative universal quantifier. For example, the universally quantified expression in (90) satisfies negative witnessing with respect to the subformula $\phi(x)$, and the quantified expression in (91) satisfies negative witnessing with respect to both $\phi(x)$ and $\psi(x)$:

(90) ...every $x(\phi(x), \psi(x))$...

(91) ...no $x(\phi(x), \psi(x))$...

But suppose that the quantifiers in (90) and (91) were NPIs, and consider what environments would license them. No matter what environment the quantifiers are embedded in, the negative witnessing condition will be trivially satisfied by the subformulas of (90) and (91) consisting of just the quantified expressions themselves. In this respect, these quantifiers are crucially different from existential quantifiers which require some broader context to be able to be part of a formula that satisfies

negative witnessing. The surprising consequence, then, is that NPI-status will not restrict the distribution of universal or negative universal quantifiers! Hence, in theory, this means that there can be negative polarity quantifiers of these sorts, but that we would never know it. In practice, though, the upshot is probably that there are no universal or negative universal quantifiers that are NPIs — for how could anyone learn that such an item was an NPI?

Another question of the same sort that we posed was why there are the particular types of environments that NPIs seem to prefer. Given that negative polarity is related to general statements, then we do not expect certain types of negative polarity environments. For example, we wouldn't expect a negative polarity existential quantifier that could occur in logical forms such as (92):

(92) at-least-two $x(\phi(x), \text{some } y(\chi(x, y), \psi(x, y)))$

(92) is not general in any sense; no reasonable variant of the negative witnessing condition will hold of it. Along the same lines, we also wanted to know what the strong and weak NPI licensing conditions have in common. In the theory I presented, the licensing conditions proposed for the two types of NPI are parameterized versions of a single NPI licensing condition. The resulting conditions differ only in how small the negative witnesses have to be. This gives, I believe, a satisfying answer to that question.

References

- Hoeksema, J. 1983. Negative polarity and the comparative. *Natural Language and Linguistic Theory* 1:403–434.
- Jackson, E. 1994. *Negative Polarity, Definites Under Quantification and General Statements*. PhD thesis, Stanford University.
- Jackson, E. 1995. Weak and strong negative polarity items: Licensing and intervention. To appear in *Linguistic Analysis*.
- Kadmon, N., and F. Landman. 1993. Any. *Linguistics and Philosophy* 16:353–422.
- Karttunen, L. 1974. Until. In *Proceedings of the Tenth Regional Meeting of the Chicago Linguistics Society*, 284–297, Chicago. Chicago Linguistics Society.
- Ladusaw, W. 1980. *Polarity Sensitivity As Inherent Scope Relations*. New York: Garland.
- Linebarger, M. 1980. *The Grammar of Negative Polarity*. PhD thesis, MIT.
- Linebarger, M. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10:325–387.

- Valencia, V. S., F. Zwarts, and T. van der Wouden. 1994. Polarity, veridicality, and temporal connectives. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the Ninth Amsterdam Colloquium*, 587–606, Amsterdam. University of Amsterdam.
- Zwarts, F. 1986. *Categoriale Grammatica en Algebraische Semantiek*. PhD thesis, University of Groningen.
- Zwarts, F. 1993. Three types of polarity. To appear in F. Hamm and E. Hinrichs (eds.), *Semantics*.