A Note on Licensing Strong NPIs

Jon Gajewski, University of Connecticut

1 Introduction

In this paper I address the licensing conditions on strong negative polarity items (NPIs). Strong negative polarity items have a more restricted distribution than so-called weak NPIs such as *any* and *ever*. The point of departure for this paper is Zwarts's (1998) elegant and influential hypothesis. Zwarts builds on Ladusaw's (1979) hypothesis that NPIs are licensed in the scope of downward entailing (DE) operators. Zwarts proposes that strong NPIs are licensed in the scope of anti-additive operators. I consider the effects of von Fintel's Strawson DE account of NPI-licensing on Zwarts's hypothesis. I show, following intuitions of Atlas (1996), Horn (1996) a.o., that Strawson anti-additive operators do not license strong NPIs. This tension between Zwarts's and von Fintel's findings is resolved with a novel theory about the licensing of strong NPIs. I propose that strong NPIs like weak NPIs are licensed by DE operators, but that, unlike weak NPIs, strong NPIs are sensitive to the interference of non-truth-conditional meaning. The resulting theory is similar in some respects to Krifka's (1995) theory.

In the rest of this section I provide the background necessary for stating the puzzle. In Section 2, I state the puzzle about the distribution of strong NPIs. In Section 3, I propose a solution to the problem, building on works of Chierchia and Krifka. In Section 4, issued raised by the new theory are discussed and tentative solutions are proposed. Section 5 briefly addresses recent alternative proposals. I conclude in Section 6.

1.1 Strong NPIs

Negative Polarity Items in English divide into (at least) two classes. The weak NPIs, such as *any* and *ever*, enjoy a wider distribution than the so-called strong NPIs. The strong NPIs that we will focus on in this paper are *in weeks*, additive *either*, and punctual *until*. For arguments in favor of the NPI status of these items see Hoeksema (1996), Rullmann (2003) and de Swart (1996)/Giannakidou (2002), respectively.[1] To get a flavor of the difference in the distribution of weak and strong NPIs observe the

[1] These authors argue against analyzing these as strong NPIs in Zwarts's sense. We will address their worries below.

acceptability of *any* and punctual *until* in the following environments.

(1) a. Bill didn't ever say anything
    b. No student ever said anything
    c. Few students ever said anything
    d. At most 5 students ever said anything
    e. *Between 5 and 10 students ever said anything
    f. *Some student ever said anything

(2) a. Bill didn't leave until his birthday
    b. No student left until his birthday
    c. ??Few students left until their birthdays
    d. *At most 5 students left until their birthdays
    e. *Between 5 and 10 students left until their birthdays
    f. *Some students left until their birthdays

While *any* is acceptable in the scope of sentential negation, negated existentials (*no student*), *few* and *at most 5*, the strong NPI *until* is only fully acceptable in the scope of sentential negation and negated existentials. There is disagreement about the acceptability of strong NPIs in the scope of *few*. Zwarts (1998) excluded it; Hoeksema (1996) and Rullmann (2003) argue that *in weeks* and *either*, respectively, are acceptable in the scope of *few*. We return to the description of the environments in which strong NPIs are acceptable later. First, let's concentrate on the environments in which weak NPIs are acceptable.

A condition that successfully divides the cases in (1) is the original description of the distribution of NPIs stated in the Fauconnier/Ladusaw Hypothesis (F/LH). The formal description of the distribution of NPIs under F/LH is the following:

(3) A (weak) NPI α is licensed only if α occurs in the scope of a downward entailing operator.

(4) Weak NPIs: *any, ever, at all…*

(5) F is Downward Entailing (DE) iff for every A, B such that $B \subseteq A$, $F(A) \Rightarrow F(B)$
    ['$\Rightarrow$' stands for cross-categorial entailment]

It is a simple matter to show, given these definitions that *not*, *no student*, *few students*, and *at most 5 students* denote DE functions and that *between five and ten* and *some* *student* do not. I leave this to the reader. One can convince oneself of DE-ness by observing intuitive entailments from sets to subsets. In other words, for example, (6) is valid but (7) is not.

(6) No student smokes
$\{x: x \text{ smokes Camels}\} \subseteq \{y: y \text{ smokes}\}$
∴No student smokes Camels

(7) Between 5 and 10 students smoke
$\{x: x \text{ smokes Camels}\} \subseteq \{y: y \text{ smokes}\}$
∴Between 5 and 10 students smoke Camels

One can see that the latter argument is invalid by considering the possibility that 5 students smoke, but only two smoke Camels. *No student* denotes a DE function, *between 5 and 10* denotes a non-monotonic function.[2]

1.2 Strawson Entailment

As successful as F/LH is, it has many well-known, long-standing problems. Von Fintel (1999) takes a significant step forward in resolving some of the problems. One kind of counterexample to F/LH is a sentence in which an NPI is licensed in an environment that is not intuitively DE.

(8) a. Only Bill ate anything.
b. Bill is sorry that he said anything
c. If Bill ate anything, then it was a hoagie

(9) a. Only Bill ate a vegetable.
#Therefore, only Bill ate kale.
b. If Bill ate something, it was a hoagie
#Therefore, if Bill ate something healthy, it was a hoagie
c. Bill is sorry he gave Mary a present
#Therefore, Bill is sorry he gave Mary a present she loved.

In (8) we see that *any* is licensed by *only*, *sorry* and in the antecedent of bare conditionals. In (9), however, we see that these environments do not intuitively license

[2] The alert reader will remember that while *between five and ten students* does not license NPIs, *exactly five students*, also non-monotonic, can. I will have nothing intelligent to say about this well-known problem case.

inferences from sets to subsets. Given the F/LH, this is troubling since we expect DE operators to license such inferences.

Von Fintel (1999) argues that in (8)a-d, it is a presupposition of the licenser that interferes with the intuitive DE-ness. He neutralizes the interference of presuppositions by redefining the notion of DE-ness relevant to NPI-licensing. Take the example of *only*. Its truth conditions are DE with respect to the predicate P:

(10) [Only a] P is defined only if **a** $\in$ P
when defined [Only **a**] P is True iff <u>There is no x≠**a** such that x $\in$ P</u>

This DE-ness is masked, however, by *only*'s presupposition. Whereas the truth conditions say that no one who is not **a** has property P, the presupposition states that **a** has property P. If no one who is not **a** is a member of P, then no one who is not **a** is a member of any subset of P. However, **a**'s being a member of P does not imply **a** is a member of every subset of P. Von Fintel redefines DE-ness as below, stipulating that to test DE-ness we must first take for granted that the presupposition of the conclusion of the set-to-subset argument is satisfied. Notice that Strawson-DE-ness is a weaker notion than DE-ness.

(11) <u>Strawson Downward Entailingness</u>
A function f of type <s,t> is Strawson-DE
iff for all x, y of type s such that x $\Rightarrow$ y and f(x) is defined : f(y) $\Rightarrow$ f(x).

Von Fintel chose to define Strawson-DE directly rather than defining Strawson Entailment and then defining directional entailment in terms of Strawson Entailment. There seems to be no reason not to, so let's do so now (following Herdan & Sharvit 2006 a.o.).

(12) <u>Cross-Categorial Strawson Entailment ($\Rightarrow_S$)</u>
a. For p, q of type t: p $\Rightarrow_S$ q iff p = False or q = True.
b. For f, g of type <s, t>: f $\Rightarrow_S$ g iff for all x of type s such that g(x) is defined: f(x) $\Rightarrow_S$ g(x).

This definition will allow us to use Strawson Entailment to define other notions relevant to NPI-licensing. This will become relevant below.

Given these definitions, we see now that *only* does come out Strawson-DE. To see this notice that the argument in (13) is intuitively valid.

(13) Only Bill ate a vegetable
Bill ate kale
{x: x is kale} ⊆ {y: y is a vegetable}
Therefore, Only Bill ate kale.

Von Fintel argues at great length that appropriate analysis of the other constructions exhibited in (8), paying special attention to the division between truth conditions and presupposition, yields meanings that satisfy the definition of Strawson-DE. Fintel's specific proposals are discussed in Appendix 1. Before doing that, let's now return to a discussion of the difference in distribution between weak and strong NPIs.

1.3 Anti-additivity

As shown above, strong NPIs have a more limited distribution than weak NPIs. While weak NPIs are clearly licensed by *few students* and *at most five students*, strong NPIs are not. They are clearly licensed only by *not* and negated existentials (*no student, never*) - limiting our view to the sentences in (2), repeated below. *At most 5 students* denotes a DE function but does not license strong NPIs. So it appears that DE-ness is not sufficient to license strong NPIs.

(2) a. Bill didn't leave until his birthday
b. No student left until his birthday
c. ??Few students left until their birthdays
d. *At most 5 students left until their birthdays
e. *Between 5 and 10 students left until their birthdays
f. *Some students left until their birthdays

The most influential account of the distribution of strong NPIs comes from Zwarts (1998), see also van der Wouden (1997). Zwarts suggests that strong NPIs require a logical property stronger than DE-ness in their licensers. Specifically, Zwarts argues that strong NPIs require anti-additive licensers. The logical property of anti-additivty (AA) is defined in (16). Note that being AA implies being DE, (17).

(14) A strong NPI α is licensed only if α occurs in the scope of an anti-additive operator.

(15) Strong NPIs: additive *either*, *in weeks*, punctual *until*

(16) F is Anti-Additive (AA) iff $F(A \vee B) \Leftrightarrow F(A) \wedge F(B)$

(17) <u>AA implies DE:</u>
(i) Suppose F is AA.
(ii) And suppose that $B \Rightarrow A$.
(iii) Then $A \vee B = A$, given (ii) [justify]
(iv) And $F(A \vee B) \Rightarrow F(A) \wedge F(B)$, given (i) and (16)
(v) But, this means $F(A) \Rightarrow F(A) \wedge F(B)$, given (iii)
(iv) and furthermore $F(A) \Rightarrow F(B)$, implied by (v)
<u>Therefore</u>, F is DE

So, anti-additivity picks out a subset of the licensers picked out by downward entailingness. The intuitive test for anti-additivity is the equivalence of wide scope conjunction with narrow scope disjunction:

(18) No student smokes or drinks $\Leftrightarrow$ No student smokes and no student drinks

(19) Few students smoke or drink $\Leftrightarrow$ Few students smoke and few students drink

No student is AA because the equivalence in (18) is valid. *Few students* is not AA because the equivalence in (19) does not hold intuitively: it could be that few students drink and few students smoke, but when you put them together and count them you get more than few. Also not AA are quantifiers such as *not every* and *At most five*. I leave confirmation of this to the reader.

AA vs. DE seems to correctly describe the differences in distribution between weak NPIs and strong NPIs. The former require a DE licenser, the latter an AA licenser. *No doctor* is AA; *At most 5 doctors* is DE but not AA.

(20) a. No doctor has seen anyone.
b. At most five doctors have seen anyone.

(21) a. No doctor has seen Mary in weeks.
b. *At most five doctors have seen Mary in weeks.

As I mentioned above, the status of *few* in this classification of licensers is controversial. While the strong NPIs listed in (15) seem to follow the Zwarts classification quite well, it has been pointed out that they sometimes tolerate *few* as a licenser.

(22) He was one of the few dogs I'd met in years that I really liked.
(Sue Grafton, A is for Alibi, Hoeksema 1996)
(23) Few Americans have ever been to Spain. Few Canadians have either.
(Rullman 2003, p.345)
(24) He invited few people$_i$ until he knew she liked them$_i$.
(de Swart 1996)

Other expressions that license strong NPIs but fail the test for AA are *hardly any/ever* and *little*. We return to this problem for Zwarts below in section 4.2.

2 AA+ Strawson Entailment: too weak!

A natural question to ask at this point is what implications these amendments to F/LH have for each other. Von Fintel attempts to resolve the problem of DE's necessity for licensing NPIs by redefining entailment. Zwarts attempts to resolve DE's insufficiency for licensing strong NPIs by replacing DE with AA in the licensing condition of those items. What happens when solutions from opposite ends of the DE problem meet?

Entailment is a fundamental relation between expressions (or their denotations). We expect a change in its definition to have far-reaching and systematic consequences in the grammar. For example, if von Fintel is right about weak NPIs, then perhaps Strawson entailment should be a primitive of NPI-licensing. That is, perhaps, anti-additivity should also be defined in terms of Strawson entailment.

(25) F is Strawson Anti-Additive (SAA) iff $F(A \vee B) \Leftrightarrow_S F(A) \wedge F(B)$[3]

Now let's look at strong NPIs under [merely Strawson] DE operators (cf. Atlas 1996, Horn 1996, Nathan 1999, Gajewski 2005, Giannakidou 2006):

---

[3] $P \Leftrightarrow_S Q$ if an only if $P \Rightarrow_S Q$ and $Q \Rightarrow_S P$.

(26) a. Only John has ever seen anyone.
b. *Only John has seen Mary in weeks
c. *Only John likes PANcakes, either. (Nathan 1999)
d. *Only John arrived until his birthday.

(27) a. If Bill has ever seen anyone, he is keeping it a secret.
b. *If Bill has seen Mary in weeks, he is keeping it a secret.
c. *If Bill likes PANcakes, either, he is keeping it a secret.
d. *If Bill arrived until Friday, he is keeping it a secret.

(28) a. Mary is sorry that she ever talked to anyone.
b. *Mary is sorry that she has talked to Bill in weeks.
c. *Mary is sorry that she likes PANcakes, either.
d. *Mary is sorry that she arrived until Friday.

(27) I have never gone to Amsterdam. *If I go to BRUSSELS either, I will buy you some chocolates. (Rullmann 2003)

(28) I didn't go to Spain. *I regret that I went to Portugal, either. (Rullmann 2003)

This is puzzling. Suppose we do define anti-additivity in terms of Strawson Entailment. Then, *only* turns out to be SAA, cf. (29) and (30) (see also Rullmann 2003 fn. 29, Gajewski 2005). And yet, strong NPIs are not licensed under these SAA operators.[4]

(29) $\underline{\textit{Only-Bill}(A \vee B) \Rightarrow_S \textit{Only-Bill}(A) \wedge \textit{Only-Bill}(B)}$
No one ≠ Bill drinks or smokes
Bill drinks or smokes
Bill drinks and Bill smokes
Therefore, No one≠Bill drinks and No one≠Bill smokes

(30) $\underline{\textit{Only-Bill}(A) \wedge \textit{Only-Bill}(B) \Rightarrow_S \textit{Only-Bill}(A \vee B)}$
No one ≠ Bill drinks and No one ≠ Bill smokes
Bill drinks and Bill smokes
Bill drinks or smokes

[4] Alonso Ovalle and Guerzoni (2002) make a similar observation about n-words in Italian and Spanish.

Therefore, No one ≠ Bill drinks or smokes

In fact, all the operators that von Fintel defines as Strawson DE come out Strawson AA. It is a routine, if tedious matter to show this (see Appendix 1).

So, a natural extension of von Fintel's notion of Strawson entailment to the licensing of strong NPIs yields incorrect results. Strawson anti-additivity is too weak to account for the distribution of strong NPIs.

To sum up, Zwarts (1998) and von Fintel (1999) have given us two dimensions to consider in the statement of licensing conditions: standard entailment vs. Strawson entailment and downward entailment vs. anti-additivity.

(31)

	Entailment	S-Entailment
DE	???	Weak NPIs
AA	strong NPIs	???

We have polarity items sensitive to AA+standard entailment and DE+Strawson entailment. To my knowledge, there are no classes of NPIs sensitive to the other two possible classes of licensers. Now we would like to know why this is so. What do anti-additivity and standard entailment have to do with each other? Why should the two go together?

This classification is an unattractive way to describe the difference between weak and strong NPIs. We have identified two independent parameters and we have a different setting for each for the two classes of NPIs. Weak = [Strawson, DE] and Strong = [Standard, AA]. A better theory would tie this two-way distinction to the setting of one binary parameter.

3. An Alternative: Sensitivity to non-truth-conditional meaning

I would like to suggest an answer to the puzzle. This answer will explain the conjunction of anti-additivity and standard entailment in a natural way and provide a framework for formulating a more perspicuous formal distinction between weak and strong NPIs.

The problem for licensers like *only* is that a non-truth-conditional aspect of meaning interferes with downward entailment. What I would like to suggest is that non-AA licensers also give rise to a non-truth-conditional implication that interferes with downward entailment.

The main idea is that weak NPIs are licensed by DE-ness defined by standard entailment when we limit our perspective to truth-conditional meaning. In other words, weak NPIs are only sensitive to truth-conditional meaning. This in fact has always been the most obvious approach to the problem presented by *only* and other presuppositional operators. Indeed, it was proposed by Ladusaw in his account of NPI-licensing under implicatives (Ladusaw 1979, p.160 ff.):

> John **failed** to buy any shirt.
> a. John **failed** to <u>buy a shirt</u>
> b. ?? John **failed** to <u>buy a red shirt</u>
>
> Since entailment depends only upon truth-conditional meaning, it will be true that (a) entails (b), even though that intuition is confused by the fact that (b) implicates something that is not implicated or entailed by (a). (a) implicates that John tried or was expected to buy a shirt, but (b) implicates that he tried to buy a red shirt. The implicature is irrelevant to the question of whether (a) entails (b).

Von Fintel justly criticizes Ladusaw for the violence he does to our *intuitive* notion of entailment. Von Fintel says,

> It just isn't good methodology to base a semantic theory on judgments about the truth of a sentence in a situation where it would be misleading and inappropriate to assert the sentence.

While this is certainly true, I do not see why it ought to preclude applying the formal relation of entailment to the truth-conditional meanings of expressions in language. Any theory that makes a division between presuppositions and truth conditions must still state some truth conditions. Our evidence about the truth conditions may be indirect because of the interference of presupposition, but it is evidence nonetheless.

That said, we cannot whole-heartedly agree with Ladusaw, either. Ladusaw asks us to just ignore non-truth-conditional meaning to assess entailments. We know, however, that we cannot do this in all cases. For weak NPIs, yes. For strong NPIs, on the other hand, we saw that the presence of presuppositions ruins the licensing capabilities of truth-conditionally anti-additive functions. What I would like to suggest is that strong NPIs are also licensed by DE-ness defined in terms of standard entailment, but the meanings that must satisfy the

DE condition are meanings that incorporate non-truth-conditional meaning.[5]

This immediately rules out (merely) Strawson DE operators as licensers of strong NPIs. We already know that if no special provisions are made (like Strawson entailment) these operators, like *only*, are not DE in the strict sense.

(32) Only Bill kicked a dog
#Therefore, only Bill kicked a terrier.

But it seems that we have only traded Strawson entailment for truth-conditions-only entailment. This does not address our criticism that weak and strong NPIs must be distinguished by different settings of two independent parameters.

To resolve this, I suggest that we can rule out non-AA operators as licenser of strong NPIs on grounds similar to those we have used to rule out (merely) Strawson DE operators as licensers of strong NPIs. Namely, we can rule them out on the grounds that they introduce a non-truth-conditional implication – a scalar implicature – that interferes with strict DE-ness. This proposal is spelled out in detail in the next section. I begin by surveying the ideas of Chierchia and Krifka that will underpin our account.

3.1 Chierchia (2004) on Intervention Effects

Inspiration for my analysis comes from two places. The first inspiration is Chierchia's (2004) analysis of interventions effects. The second inspiration is Krifka's 1995 suggested generalization about strong NPI licensers.

Chierchia argues that intervention effects in NPI licensing can be analyzed as the interference of an implicature with the licensing of an NPI. More specifically, Chierchia argues that the class of interveners can be identified as the class of expressions that do not sit at the weak end of a Horn scale. Whenever such an item falls between a DE expression and an NPI, a quantity implicature is generated since a stronger statement could have been made by replacing the item with the weak endpoint of the scale.

[5] I intend to include here, non-truth-conditional meaning that is generalized in Grice's sense, as opposed to particularized.

This implicature is added to the meaning of the minimal constituent containing the DE operator. The addition of the implicature interferes with downward entailment. For example, *every* NP, which is not a weak scalar endpoint, introduces an implicature when in the scope of a DE operator. This implicature makes *every* NP an intervener for NPI-licensing. (Assume that *any* is an existential quantifier like *some*:)

(33) The quantifier *everyone* intervenes for licensing of *anything*:
 *Bill didn't give everyone anything.
 NOT Bill gave EVERYone ANYthing

The reason is that a weaker member of *every* NP's scale could have been used, yielding a globally stronger statement, cf. (34).

(34) Stronger Alternative to (33):
 NOT Bill gave SOMEone ANYthing

Hence, by standard scalar Gricean reasoning, this stronger alternative is taken to be false.

(35) Implicature (negation of stronger alternative):
 NOT NOT Bill gave SOMEone ANYthing
 "Bill gave someone something"

The 'strong meaning' for the sentence (33), then, is the conjunction of its plain meaning and the negation of the stronger alternative, (35):

(36) Strong Meaning:
 (NOT Bill gave EVERYone ANYthing) AND (Bill gave SOMEONE ANYthing)

Notice now that this strong meaning does not create a DE environment in the position of the NPI. For example the set-to-subset inference from things (36) to books (37) is not valid.

(37) Strong Meaning not DE since (36) does not entail, for example:
 (NOT Bill gave EVERYone ANYbook) AND (Bill gave SOMEONE ANYbook)

In this way, the presence of a non-weak scalar endpoint between an NPI and a DE operator can block licensing by destroying DE-ness.

So, Chierchia argues that implicatures can insinuate themselves into the licensing of NPIs. At first this might seem odd, since many intervention effects are irrevocable, whereas implicatures are by definition defeasible. Chierchia urges us to ignore the fact that context can override an implicature and simply focus on the role implicatures play in the compositional interpretation of sentences. In other words, the licensing conditions of NPI make reference to the recursive contributions of implicatures, even if that contribution has a will o' the wisp character in context.

This is a beautiful story about intervention effects, but it must be carefully formulated, as Chierchia notes. If the DE licenser <u>itself</u> is not at the strong end of its Horn scale, it will itself contribute an implicature that would destroy downward entailment (Chierchia credits Dominique Sportiche with this observation). Consider the case of *few* which is weaker than *no* on the DE scale of determiners. Used in an appropriate context, *few* gives rise to the implicature that *no* would not hold in its place. That is, it implies that something falls in the intersection of its restrictor and scope. This implicature interferes with downward entailment in the same way that the positive presupposition of *only* does: From the propositions (i) that few students read anything and (ii) that some students read something, we cannot conclude that some students read some book (the implicature of *few students read any book*) – they may all have read pamphlets.

(38) FEW students read ANYthing

(39) Stronger Alternative:
NO students read ANYthing

(40) Implicature:
NOT(NO students read ANYthing)
=SOME student read ANYthing
"some student read something"

(41) Strong Meaning:
(FEW students read ANYthing) AND (SOME student read ANYthing)

(42) Strong meaning not DE since (38) does not entail:
(FEW students read ANY book) AND (SOME student read ANY book)

Chierchia makes the necessary adjustments to his theory to prevent the implicatures of the licensers from interfering with weak NPI licensing. Specifically, he divides implicatures into direct and indirect implicatures. Within Chierchia's system, these two kind of implicatures are introduced by two different rules: indirect implicatures are introduced by Strong Functional Application, which is sensitive to DE-ness, and direct implicatures are introduced by Krifka's Rule. The implicatures of non-weak endpoints in the scope of DE operators are indirect implicatures. Implicatures of non-strong endpoints (like *few*) are direct implicatures. I refer readers to Chierchia's work for further details.

We will make use of Chierchia's direct implicatures in our formulation of the licensing principle for strong NPIs.

3.2 Krifka (1995) on Strong NPIs

This leads us directly to our second source of inspiration: Krifka's (1995) restatement of the distribution of strong NPIs. Krifka suggests that anti-additivity is not in fact the relevant property for licensing strong NPIs.[6] He argues that instead the relevant property is being at the end of a DE scale.[7] Krifka says this because he believes that strong NPIs are emphatic and that emphatic items need to be in extreme environments, such as the scope of an operator at a (negative) scalar endpoint. This is somewhat vague as an explanation, but the generalization has interesting consequences.

Notice what is true if Krifka is right. If the licensers of strong NPIs are all strong scalar endpoints, then they do not introduce quantity implicatures (locally). So, for

---

[6] Be aware that what Krifka means by strong NPI does not coincide exactly with what Zwarts means. Krifka includes stressed *ANY* in the class of strong NPIs. Stressed *ANY* has a broader distribution than Zwarts's strong NPIs. For example, it can occur in the ostensibly upward entailing complement of *glad*:

(i) I'm glad we got ANY tickets.

[7] See Matsumoto (1995) for an argument that scales must be uniform in monotonicity.

example, the use of *few* invites the implicature *some*, but the use of *no* invites no implicature.

(43) CLAIM: strong NPI licensers = strong endpoints of DE scales (Krifka 1995)

(44) Strong NPI licensers: *not*, *no* NP, *never*

<NOT>
<*NO* NP, *FEW* NP, *NOT EVERY* NP>
<NEVER, RARELY, NOT ALWAYS>

Now we see that (merely) Strawson AA and merely DE operators form a natural class. They both introduce a non-truth-conditional aspect of meaning that intereferes with judgments of monotonicity.

What I would like to suggest then, to put it in simplified form, is that weak NPIs are licensed by downward entailment in the truth-conditional meaning (and can ignore non-truth-conditional meaning), whereas strong NPIs can only be licensed by downward entailingness that is preserved when conventional non-truth-conditional meaning is taken into consideration. To accommodate Chierchia's account of intervention, this statement would need to be modified to allow weak NPIs to be sensitive to some kinds of non-truth-conditional meaning (in Chierchia's terms *indirect* but not *direct* implicatures).

Notice how this is superior to an account that simply throws Zwarts and von Fintel together. Instead of differentiating weak NPIs from strong NPIs with two independent features, we can differentiate them based on one feature: sensitivity to the (conventional) non-truth-conditional meaning of their licensers.

3.3 Stating the Licensing Conditions

I will give here a fairly conservative formalization of the licensing principles of NPIs suggested by our observations. I take from Chierchia a formal approach to assigning strengthened meanings to constituents, where the strenghthened meaning is a constituent's plain meaning (which includes presuppositions) plus its scalar implicatures. The strenghthened meanings will be the meanings tested for DE-ness for the licensing conditions of strong NPIs. I take from von Fintel the formulation of Strawson-entailment and use it to move the presuppositions of plain, unstrengthened meanings out of the way to assess DE-ness for the licensing of weak NPIs.

The key elements of the proposal are (i) its use of DE-ness for licensing both weak and strong NPIs and (ii) its grouping presuppositions and scalar implicatures together as non-truth-conditional aspects of meaning that function together in the statement of licensing principles.

3.3.1 An implementation

I treat presuppositions as domain conditions on functions in the semantics. So, to neutralize their effect on weak NPIs I use von Fintel's Strawson entailment, repeated below in (46). So, our treatment of weak NPIs is identical to von Fintel's (1999).

(45) $[\![\textbf{only}]\!] = \lambda x.\lambda P\text{: } P(x)=1.\ \forall y[y \neq x \rightarrow P(y)=0]$

(46) <u>Cross-Categorial Strawson Entailment ($\Rightarrow_S$)</u>

a. For p, q of type t: $p \Rightarrow_S q$ iff p = False or q = True.

b. For f, g of type <s, t>: $f \Rightarrow_S g$ iff for all x of type s such that g(x) is defined: $f(x) \Rightarrow_S g(x)$.

For strong NPIs, on the other hand, I state the licensing condition in terms of standard entailment applied to meanings enriched with scalar implicatures. I represent such enriched meanings with a covert exhaustivity operator EXH (following Fox 2003/2007). EXH operates on the meaning of a constituent against a background of alternatives. The set of alternatives is generated by Chierchia's (2004) ALT function. This function delivers the alternatives 'induced solely by the by the last scalar element in the tree (i.e. the highest or topmost one)' (Chierchia 2004 p. 17). So, the alternatives to a DP will be those induced by the alternatives to the determiner.

In (47), I give examples of the alternatives to determiners, which are derived from their Horn scales. In (48), I show how the alternatives to determiners in (47) project into alternatives to the DPs they head, see Rooth (1985) for compositional methods of composing sets of alternatives.

(47) a. $[\![\textbf{few}]\!]^{ALT} = \{[\![\textbf{no}]\!], [\![\textbf{few}]\!], [\![\textbf{not many}]\!], [\![\textbf{not every}]\!]\}$

b. $[\![\textbf{fewer than 3}]\!]^{ALT} = \{ \ldots [\![\textbf{fewer than 2}]\!], [\![\textbf{fewer than 3}]\!], \ldots [\![\textbf{fewer than n}]\!] \ldots \}$

(48) a. $[\![\alpha]\!]^{ALT} = [\![\beta]\!]^{ALT}([\![\gamma]\!])$

b. $[\![\textbf{few students}]\!]^{ALT} = \{[\![\textbf{no student}]\!], \ldots [\![\textbf{not every student}]\!] \}$

In (49), I give a rule for the implicature-enriched meaning of a generalized quantifier. This schema could easily be made cross-categorial in the spirit of Rooth's (1985) approach to *only*.

(49) $[\![\text{EXH } Q]\!]^{w}(C) = \lambda P_{<e,t>}.[\![Q]\!]^{w}(P)=1 \ \& \ \forall Q' \in C[Q'(w)(P)=1 \rightarrow \forall w'[\ [\![Q]\!]^{w'}(P)=1 \rightarrow Q'(w')(P)=1]\ ]$

Now we can give a formal statement of the licensing conditions for weak and strong NPIs in English. As weak NPIs are only sensitive to truth conditions, their licensing condition only inspects the Strawson-entailingness of its licenser. Since strong NPIs are sensitive to both truth-conditional and non-truth-conditional meaning, their licensing condition inspects the (plain) entailingness of their licenser's enriched meaning.

Licensing Principles:

(50) A weak NPI α is licensed only if it occurs in the scope of β, where $[\![\beta]\!]$ is SDE

(51) A strong NPI α is licensed only if it occurs in the scope of β, where $[\![\text{EXH } \beta]\!]([\![\beta]\!]^{ALT})$ is DE

Because *no* sits at the end of its scale, its enriched meaning is the same as its plain meaning (52), which is DE. So, we expect it to license strong NPIs. A determiner like *not every*, on the other hand, does not sit at a scalar endpoint and thus gives rise to implicatures in neutral contexts. This leads to an enriched meaning like that in (53), which, because of its UE component (some students), is not DE.

(52) $[\![\text{EXH } \textbf{no students}]\!]([\![\textbf{no students}]\!]^{ALT}) = [\![\textbf{no students}]\!]$

(53) $[\![\text{EXH } \textbf{not every student}]\!]([\![\textbf{not every student}]\!]^{ALT}) = [\![\textbf{some student but not every student}]\!]$

An analysis that separates truth conditions and presuppositions into separate dimensions of meaning

(Karttunen & Peters 1979; Horn 2002, to appear) would perhaps be more elegant by allowing both licensing conditions to be stated in terms of standard entailment. I leave such a radical project for future research.

3.3.2 Implications for Intervention

Note that the licensing condition on weak NPIs does not make use of strong meanings. Consequently, we lose Chierchia's account of intervention. Recall, though, that Chierchia (2004) needs to ignore certain implicatures in the licensing of weak NPIs. DE expressions like *not many* have implicatures of their own.

(54) Not many students left
Strong meaning: Not many students left and some students left.

To prevent these implicatures from interfering with licensing, Chierchia draws a distinction between direct and indirect implicatures. Indirect implicatures are implicatures introduced by reversal at DE nodes. Chierchia claims that only these interfere with NPI licensing. We might attempt to incorporate this distinction into our story:

(55) Entailments, Indirect Implicatures vs.
Direct Implicatures, Presupposition

There does not, however, appear to be any independent support for such a division. Another option would be to adopt a different, perhaps syntactic, view of intervention. Some have argued that intervention in NPI licensing should be seen as part of a broader phenomenon (e.g., Beck effects). See Guerzoni (2006) for a recent view of this kind. It's not clear how Chierchia's proposal could extend to other cases of intervention.

Doubt is cast on such a syntactic account by Homer's (2009) observation that presuppositional items also induce intervention effects in NPI licensing. For example, Homer observes that the presence of the presuppositional particle *too* in (56)a interferes with the licensing of *anything* by matrix negation.

(56) Context: Mary read some interesting book.
a.*I don't think $[\text{John}]_F$ read anything interesting too.
b.I don't think $[\text{John}]_F$ read something interesting too.

c. Presupposition of (3b): Somebody other than John read something interesting.

If this generalization holds up, it pushes us to the following conclusion. Implicatures and presupposition introduced between an NPI and its licenser always trigger interventions effects; implicatures and presuppositions introduced by an NPI licenser interfere with the licensing of strong NPIs, but not with the licensing of weak NPIs. At this point, I do not have an explanation of why this is so. Ultimately, the answer should be sought in the semantics of the NPIs themselves. The licensing conditions stated in this paper should be understood as standing proxy for a deeper explanation of why the expressions discussed are polarity sensitive in the first place. If we understand how the sensitivity of *either*, for example, derived from its semantics, we would be better situated to answer the question of why it is sensitive to the non-truthconditional aspects of its licenser's meaning while *any* is not.

I leave a reconciliation with Chierchia's (2004) and Homer's (2009) results for future research. In the next section, we turn to issues raised by our new proposal.

4 Potential Challenges

The statement of our new theory – according to which strong NPIs are licensed by DE-ness when non-truth-conditional meaning is taken into account – raises the possibility of treating certain licenser in new ways. We need to be sure that expressions that do not license strong NPIs are still predicted not to. This is the case of comparative quantifiers discussed in section 4.1. We may also be able to account now for expressions that do license strong NPIs but are not AA. This is the case of *few* discussed in section 4.2. Finally in Section 4.3, we suggest a new perspective on a problem that plagues semantics accounts of licensing: the semantic equivalence of *no* and (*exactly*) *zero*.

4.1 Comparative quantifiers

What our theory really says now is that strong NPI licensers are DE operators that introduce neither presuppositions nor (local) quantity implicatures.[8] I have

[8] Note that this statement is not completely accurate. Not just any presupposition associated with a function will

suggested identifying this set of licensers with the set of strong endpoints on DE scales. I may have suggested this identification too quickly. Might there be DE operators that are not scalar endpoints but also do not introduce implicatures? As far as I can tell there are two possibilities. The first possibility: the operator is not a member of a scale (see Chierchia's 2004 analysis of *if* clauses). The second possibility: the operator qualifies as a member of a scale, but does not give rise to an implicature. The second possibility may well be attested. Krifka (1999) and Fox & Hackl (2006) (F&H) suggest that some scalar expressions do not give rise to implicatures.

F&H suggest that the reason the implicatures do not arise is that admitting such implicatures would lead to contradiction. We need not get deep into technical details, but let's get the flavor of their analysis. The case they focus on is that of comparative quantifiers, such as *more than n*. While (57) gives rise to the implicature that Mary didn't eat four éclairs, the roughly equivalent (58) does not.

(57) Mary ate three éclairs
     Implicature: Mary didn't eat four éclairs.

(58) Mary ate more than two éclairs.
     #Implicature: Mary didn't eat four éclairs.

They suggest that this judgment is supported by the following contrast (in their view implicatures are introduced by an operator with semantics similar to *only*).

(59) a. Mary only ate [three]$_F$ éclairs.
     b. #Mary only ate more than [two]$_F$ éclairs.

F&H suggest that no implicature is derived from *more than n* because running the usual implicature generation mechanism leads to contradiction. Such a contradictory strengthened meaning is useless and is, therefore, discarded. How does the contradiction arise? F&H suggest that the relevant alternatives to *more than n* are *more than m* for some number m. Furthermore, they suggest that the default output of

---

prevent it from licensing strong NPIs. For example, if there is an existence presupposition associated with *no student* that will <u>not</u> stand in that way of the functions DE-ness, licensing strong NPIs in its scope.

scalar implicature is the negation of all stronger alternatives.

(60) Mary ate more than 3 éclairs.
Mary ate more than m éclairs.
NOT(Mary ate more than m éclairs), for all m > 3

The negation of all stronger such alternatives is incompatible with the truth conditions of the original statement under the assumption that the domain of number alternatives is dense. In fact, F&H suggest that in natural language all measurement domains are dense:

(61) Universal Density: $\forall d1,d2\ [d1>d2 \rightarrow \exists d3(d1>d3>d2)]$

Saying that Mary ate more than 3 éclairs implies that she ate n éclairs for some n > 3. If we then say however that *Mary ate more than m éclairs* is false for all m > 3, we run into a contradiction. Because the domain of numbers is dense there is a number o between 3 and n. But *Mary ate more than o éclairs* must be false since o > 3. That implies that Mary didn't eat more than o éclairs. But n > o.

F&H also discuss the case of negated numerals:

(62) Bill didn't smoke 30 cigarettes.
#Implicature: Bill smoked 29 cigarettes.

According to F&H, *not n* is interpreted as *fewer than n*. Their account of the lack of implicatures for *more than n* extends directly to *fewer than n*, so that no impicature is generated in (62).

Krifka (1999), on the other hand, explicitly claims that negative comparative quantifiers, such as *fewer than/at most* n NP, do not give rise to implicatures. If this is true, it is incompatible with the view of NPI-licensing being put forward here. *Fewer than n* is DE, if it does not introduce an implicature we predict it should license strong NPIs. This is not the case (though see Krifka 1995[9]):

---

[9] Krifka suggests (i) can be good (he takes stressed *at all* to be a strong NPI):

(i) Fewer than three students talked AT ALL.

(63) *Fewer than 3 students have visited in weeks.

One way out for us is to suggest that, even though comparative quantifiers do not introduce the same implicatures as bare numerals, they do introduce some implicature.

For example, (62) clearly implicates that Bill smoked some cigarettes, if not 29. It may be that (64) also introduces such an existential implicature.

(64) Fewer than ten students attended the colloquium.

To my ear, (64) introduces the implicature that some students attended the colloquium. Of course, in some contexts this implicature, like any other, can be canceled. Krifka (1999) admits that sentences like (64) typically give rise to an existential inference. He supports this with evidence from anaphora. In particular, he claims that it is (almost) natural to follow the statement (65)a with reference in (65)b to the entity introduced by the existential contribution of *fewer than three students*.

(65) a. Fewer than 3 students left early.
     b. ?And they only left because they felt ill.

Krifka, however, takes this existential inference to be a presupposition. I am not convinced this is a presupposition. Consider the following tests. First, we apply von Fintel's (2004) *hey wait a minute* test. It sounds quite odd to object to (65)a with (66)a. This suggests that negative comparative quantifiers do not carry an existential presupposition. If existence were being taken for granted such an objection should be felicitous.

(66) #Hey wait a minute! I had no idea some students left early.

Furthermore, the existential inference does not project like a presupposition. For example, (67) ought to carry the presupposition that Mary believes some students will walk out on her talk (cf. Heim 1992). This does not appear to be the case.

(67) Mary wants fewer than 3 students to walk out on her talk.

As I have said above, I believe that the existence inference is an implicature. But if Fox and Hackl are correct, (68) holds.

(68) [[EXH fewer than 3 students]] (P) = $\perp$

If all implicatures are introduced by EXH, this suggests that *fewer than 3 students* cannot introduce an implicature. I suggest that if the strengthening operator EXH produces inconsistency, a weaker one (W-EXH) steps in to generate an existential implicature:[10]

(69) [[W-EXH Q]](C)=(C is the set of alternatives to Q)
$\lambda w.\lambda P_{<e,t>}$. [[Q]] w(P)=1 & $\exists Q' \in C$[Q'(w)(P)=0]

In this way, we can say that negative comparative quantifiers, such as *fewer than* n NP, give rise to existential implicatures. This existential implicature interferes with DE-ness in the same way as *only*'s presupposition. Hence we predict that negative comparative quantifiers will not license strong NPIs, though they do license weak NPIs. In the next section, we return to the difficult case of licensers like *few*, *little*, and *hardly any*/*ever*.

4.2 *Few*: a potential advantage?

It is well known that in certain contexts, non-AA functions can license strong NPIs. Zwarts's approach makes no allowance for this. For example, *few* is DE but not AA.

(70) He was one of the few dogs I'd met in years that I really liked.
(Sue Grafton, A is for Alibi, Hoeksema ms.)
(71) Few Americans have ever been to Spain. Few Canadians have either.
(Rullman (2003), p.345)
(72) He invited few people$_i$ until he knew she liked them$_i$.

[10] Danny Fox (p.c.) asks what would happen if instead we had only one EXH, but the semantics of EXH used Innocent Exclusion (IE) as suggested in Fox (2007). Interestingly, a contradiction still arises in the comparative quantifier cases even if EXH is defined with IE. This contradicts Fox (2008), which suggests that IE EXH is 'contradiction free'. See Gajewski (2009) for discussion.

(de Swart 1996)

Our analysis, on the other hand, can make allowances for such cases. Chierchia 2004 argues that items near the end of a scale can in some contexts behave as if they were the scalar endpoints. As an example, he gives the case of *many* not causing intervention effects. If *many* is on a scale with *some*, it should trigger an implicature in the scope of a DE operator. Yet, intervention effects by *many* are context dependent.

(73) I typically don't have many students with any background in linguistics.
(74) SOME <MANY, EVERY>

Chierchia suggests that when (73) is acceptable, it is because *many* sits at the endpoint of the scale <u>in context</u>, as in (74). According to Chierchia, a universal quantifier can never count as the weak endpoint of a scale in any context and, thus, always intervenes. This follows from Chierchia's axiom on scale structure that says that a scale must always contain at least two items.

This suggests that negative scales can be truncated in context as well. So *few*, being just above *no*, may serve as the strong endpoint of the negative scale. In such a context, no implicature is generated by *few*. Chierchia 2004 justifies leaving *some* off of *many*'s scale in the following way: "What enables us to truncate a scale at the low end […] is that small amounts may be functionally equivalent to nothing." Chierchia notes that sentence such as (75) need not carry implicatures.

(75) I typically don't have many students with any background in linguistics.

I propose to transpose Chierchia's reasoning about *not…many* to the case of *few,* cf. (76). In the same contexts that allow *some* to be left of *many*'s scale in the scope of a DE operator, let *no* be allowed to be left off *few*'s scale,(77).[11]

---

[11] Note that this violates Chierchia's (2004) scale axiom (i):

(i) In any given context where we utter a sentence S, containing a scalar term α:
if possible, α must not be the strongest element of the chosen scale.

(76) Typically, few students in my class take an interest in semantics.

(77) NO <FEW, NOT EVERY>

Consequently, strong NPIs can be licensed by *few* when context permits.

This is all a bit vague though; let me propose a precise restriction on when a negative operator can act like a strong scalar endpoint.

(78) **Condition on Truncation of negative scales:** to be able to act as a strong scalar endpoint a scalar item must be close enough to the endpoint.

I propose that to be considered "close enough" a scalar item must be Intolerant (see Löbner 1985, Horn 1989[12]). Horn 1989 uses the concept of Intolerance to identify those items that are above the midpoint of a scale. This provides an interesting precedent for our account of being near the endpoint of a scale.

(79) A function f of type <<e,t>,t> is Intolerant iff if f is not trivial[13], then for all x of type <e,t>, $f(x)=0$ or $f(\neg x)=0$.

(80) A function f is trivial iff for all x, $f(x)=1$ or for all x, $f(x)=0$.

On its proportional reading, *few* is plausibly Intolerant. *Fewer than 4* is not; (82) can be true if I have at most six friends.

(81) a. #Few of my friends are linguists and few of them aren't.
(Horn 1989)
b. #He rarely goes to church and he rarely doesn't go.
(Horn 1989)

Chierchia discusses the case of positive scales. Perhaps, the axiom must be reversed for negative scales.

[12] See also Zwarts's (1998) discussion of the Law of Contradiction.

[13] I include this clause to bring out the inclusion relations in (83). See Appendix for proof that AA $\subseteq$ DE+Intolerant.

(82) Fewer than 4 of my friends are linguists and fewer than 4 aren't.

Thus, while *few* NP may license strong NPIs, *fewer than n* NP may never – though see section 4.3 below.

Someone unconvinced by the details of my story might still be interested in DE+Intolerant as an intermediate category of negation between DE and AA. In fact, the following inclusion relations hold.

(83) AA $\subset$ DE+Intolerant $\subset$ DE

Finally, I find additional support for the proposed condition on scale truncation (78) in the details of the NPI-licensing properties of *few*. Partee (1989) argues that *few* is ambiguous between a proportional and a cardinal reading. On its cardinal reading *few* has a meaning like (84)a; on its proportional reading (84)b.

(84) a. $[[\text{ few }]](A)(B) = 1$ iff $|A \cap B| < n$, where n is small.
b. $[[\text{ few }]](A)(B) = 1$ iff $|A \cap B| < n \cdot |A|$, where $n<1$ and n is small.

The cardinal reading is not Intolerant, but the proportional reading is Intolerant – whenever n is less than ½. So, in environments where the cardinal reading is forced, *few* should not be able to license strong NPIs. Existential *there*-sentences have been argued to force the cardinal reading of *few*. Thus we predict, correctly, that (85)b is ungrammatical.

(85) a. There were few potatoes in the pantry.
b.?*There were few in the refrigerator, either.

This shows, I believe, that cardinal *few* never licenses strong NPIs. In the next suggestion we turn to possible problems for the current proposal raised by certain numeral quantifiers.

<u>4.3 *Zero* and explicit proportions</u>

Semantic accounts of strong NPI licensing are haunted by the problem of semantic equivalence. While *no* is ostensibly semantically equivalent to *fewer than one* and *zero*, only *no* licenses strong NPIs in its scope.

(86) [[no]] = [[fewer than one]] = [[exactly zero]]

(87) a. *Fewer than one student has visited me in years.
b. *Exactly zero students have visited me in years.

My theory does no better than an AA theory here, since *fewer than one student* does not intuitively give rise to such an existential implicature. One possible response is to follow Fox and Hackl (2006) in assuming all measurement domains are dense. The system will produce an implicature like ".3 students left" but the implicature doesn't see the light of day once it confronts our world knowledge about counting students.

This is fine for *fewer than one* NP, but it will not work for (exactly) zero NP. Here is an alternative that could take care of *zero* NP. Suppose the grammar (and implicature-generating mechanism as a part of it) can't distinguish one numeral from another. The grammar knows degree domains are ordered and possibly dense but doesn't know the names of degrees. Recall that functions like [[zero students]] are intuitively DE (even AA) and do not give rise to positive implicatures, but do not license strong NPIs. In fact, there are many ways in which (*exactly*) *zero* behaves differently from *no*:

(88) Zero students left early
No/*Zero students like SEMANTICS, either.
(89) a. On no/*zero occasion(s) did he mention my help. (cf. Deprez 1999)
b. No/*Zero students but Bill came. (cf. Moltmann 1995)
c. She drank no/*zero martinis, not even weak ones. (cf. Postal 2004)
(90) ?Zero students said anything.

We could explain this if the grammar sees *zero* as just another number, like *sixty four*. Suppose that the grammar only ascribes a property to an expression containing a numeral n if it has that property on all values for n. Since (*exactly*) n is not Intolerant on all values, the grammar does not acknowledge it as such. Therefore, (*exactly*) *zero* cannot serve as the endpoint of a scale. Now we turn to a problem with the Intolerance condition.

I argued for Intolerance as a line dividing DE quantifiers that could act as endpoints from those that could not. Explicit proportionals like (91) are a problem. *Fewer than 1/3* NP is indeed Intolerant, but does not license strong NPIs.

(91) *Fewer than 1/3 of the students have visited in weeks.

Perhaps, grammar is not good at working out explicit proportions. Fox (2000) argues for a similar conclusion on the following grounds. Fox argues that wide scope is possible with respect to negation for the object quantifier in (92)a. I refer the reader to Fox (2000) for the detailed arguments. The reason this is of interest is the Fox's economy conditions only allow a quantifier to take wide scope if the reading thereby derived is distinct from the narrow scope reading. In (92)a, this is not the case.

(92) a. Rob doesn't speak more than half of the 9 languages spoken in Sydney. (Fox 2000)
b. Rob doesn't speak 5 of the 9 languages spoken in Sydney.

Fox tentatively argues that the grammar, which evaluates violations of the economy condition, does not have access to the mathematical content of words such as *half*. My own tentative conjecture, extending Fox's observation to the case of (92)b, is that the grammar cannot ascribe different grammatical properties to expressions because they contain different numeral expressions.

4.4 Summary

In this section we have dealt with several issues raised by the novel properties of the approach to licensing strong NPIs proposed in Section 3. First, we tackled the case of comparative numeral quantifiers. It has been claimed that, despite their scalar nature, these quantifiers do not give rise to scalar implicatures. This would mean that we predict, contrary to fact, that *fewer than 3* NP licenses strong NPIs. I argued that comparative quantifiers do carry implicatures, though weaker ones than might be expected. Second, we saw that the structure of the theory proposed in Section 3 gives us a new perspective on the ability of *few*, *little* and *hardly any* to license strong NPIs. The crucial notions were scale truncation (Chierchia 2004) and Intolerance (Horn 1989). Finally, we observed that the special status of numeral expressions in the grammar gives us a chance to understand the differences in licensing between *no* and *zero*; and saves us from incorrect predictions concerning the licensing properties of proportionals like *fewer than 1/3*.

In the next section I briefly comment on some recent alternative proposals concerning similar data sets.

5 Comments on Recent Alternatives

In this section I lay out and criticize two recent alternative proposals concerning the licensing of strong NPIs. The first is the antiveridicality and rescuing approach of Giannakidou (2006); the second is Levinson's (2008) approach based on 'semantic negativity.'

5.1 Giannakidou 2006

Giannakidou 2006 also addresses the distinction between strong and weak NPIs. She argues that strong NPIs are licensed by being in the scope of nonveridical, specifically antiveridical, operators, cf. (94).

(93) **DEFINITION 3.** (Non)veridicality for propositional operators
i. A propositional operator *F* is veridical iff *Fp* entails or presupposes that *p* is true in some individual's epistemic model ME(x); otherwise *F* is nonveridical.
ii. A nonveridical operator *F* is *anti*veridical iff *Fp* entails that *not p* in some individual's epistemic model: $Fp \rightarrow \neg p$ in some ME(x).

(94) *Licensing by nonveridicality*
A polarity item α will be grammatical in a sentence S iff α is in the scope of a nonveridical operator β in S.

How to extend nonveridicality to non-propositional operators is a matter that has never fully been resolved. Nor is it clear that antiveridicality picks out the right class of licensers as delineated by Zwarts 1998.

Giannakidou observes that Strawson DE operators such as *only* do not license strong NPIs. She does not endorse Strawson entailment, but rather analyzes *only* in the spirit of Atlas. This means that she assumes the truth of the prejacent is part of the truth conditions of an *only* statement.

(95) Atlas (1991, 1993): *only a P* asserts:
$\exists x \forall y[(x=y \leftrightarrow Py) \;\&\; (Py \rightarrow y=a)]$
= Exactly one individual, and no one other than a,

has the property P.
Which *entails* the positive proposition: P(a)

This means, in her terms, that *only* is veridical, so cannot license strong NPIs. *Only* of course does license weak NPIs – this was the motivation for von Fintel's Strawson DE account. Consequently, Giannakidou 2006 suggests that weak NPIs are not licensed, but rescued by a negative proposition made available by the sentence containing the weak NPI.

(76) *Rescuing by nonveridicality*
A PI α can be rescued in the scope of a veridical expression γ in a sentence S, if (a) the global context C of S makes a proposition S' available which contains a nonveridical expression β; and (b) α can be associated with β in S'.

She states: "In the case of *only*, we saw that the nonveridical proposition is an entailment of the sentence (the non-cancelable exclusive conjunct); in the case of negative emotive factives it is possibly a conventional implicature (a counterfactual containing negation)."

This account inherits all the problems of a negative implicature licensing system, like Linebarger (1987). Despite Giannakidou's claims, her system faces the problem of overgeneration that was never resolved in Linebarger's theory.

5.2 Levinson 2008

Levinson (2008) deals with a substantially similar set of facts as the present paper, though he chooses as his domain of inquiry what he calls negative polarity particles (NPPs): *either, neither, yet* and *already*. Developing a proposal of Rullmann (2003), Levinson claims that NPPs are licensed by what he calls 'semantic negativity.' More specifically, an NPP is licensed if the clause containing it is semantically negative. Semantic negativity consists of two notions: downward monotonicity and assertivity. Definitions are given below.

(96) Levinson's licensing conditions:
A clause *x* is *downward monotone* (DM) relative to *z* if the predicate position of *x* is downward monotone in *z*.
x is *assertive relative* to a z iff

x = z or x is a subclause of z, and for some assertion strength i,

$\text{ASSERT}(z) \Rightarrow_{ILL} \text{ASSERT}_i(x)$ or $\text{ASSERT}(z) \Rightarrow_{ILL} \text{ASSERT}_i(\neg x)$

(97) x is *semantically negative relative* to z iff
x = z or x is a subclause of z and
x is DM relative to z and
x is assertive relative to z

(98) x is *semantically negative* iff there exists z such that x is semantically negative relative to z.

(99) NPPs are licensed in a host clause *x* if *x* is *semantically negative*.

The problem with this approach is that it is essentially clausal. Since every matrix clause is assertive with respect to itself, licensing within a matrix clause reduces to downward monotonicity. This leaves, as Levinson notes (p. 160), no room for differentiating between DM licensers. That is, Levinson's account predicts (100) to be grammatical; the matrix clause is both DM and assertive.

(100) *At most five students like PANCAKES, either.

I hope to have shown in section 1 of this paper, that the confinement of strong NPIs, such as *either*, to a subset of DE environments is a fundamental property of their distribution and one that any account of their licensing must deal with.

6 Conclusion

The main conclusion of this paper is that both weak and strong NPIs seek DE licensers. Weak NPIs look for them in the truth conditions. Strong NPIs, on the other hand, are forced to take presuppositions and implicatures into account when assessing DE-ness. This explains observations about the distribution of strong NPIs with respect to presuppositional licensers (as in Atlas 1996, Horn 1996 a.o.), as well as Zwarts's strength hierarchy. Furthermore, we have argued that (i) DE comparative quantifiers do give rise to scalar implicatures and thus cannot license strong NPIs (ii) non-scalar endpoints, like *few*, can sometimes act as a scalar endpoint in context if they are Intolerant and can, therefore, license strong NPIs (iii) *fewer than 1/3* and *zero* may not be counterexamples to our theory, if we are correct that grammars cannot ascribe properties such as Intolerance to a phrase based on the identity of a numeral. The main issue left open for further research is how the picture of licensing advanced here jibes with observation about intervention made by Chierchia (2004) and Homer (2009).

APPENDIX 1

In this appendix, I summarize von Fintel's analyses of adversatives, conditionals and superlatives and demonstrate that all turn out to be Strawson anti-additive. First, here are some preliminary defintions:

For any set of propositions P, we define a strict partial order $<_P$:

$$\forall w', \forall w'': (w' <_P w'' \text{ iff } \forall p \in P\ (w'' \in p \rightarrow w' \in p \text{ and } \exists p \in P\ (w' \in p \ \&\ w'' \notin p)\ )$$

In plain English:

w′ is better than w″ according to P iff all propositions in P that hold in w″ also hold in w′ but some hold in w′ that do not also hold in w″.

For a given strict partial order $<_P$ on worlds, define the selection function max_P that selects the set of $<_P$-best worlds from any set X of :

$$\forall X \subseteq W: max_P(X) = \{w \in X: \sim\exists w' \in X: w' <_P w\}.$$

With these in mind, von Fintel gives the semantics of the adversative predicate *sorry*.

$[[\ sorry_i\ ]]^{f,g}(p)(a)(w)$ is defined only if

(i) $DOX(\alpha,w) \subseteq p$
(ii) $DOX(\alpha,w) \subseteq f_i(\alpha,w)$
(iii) $f_i(\alpha,w) \cap p \neq \varnothing$
(iv) $f_i(\alpha,w) - p \neq \varnothing$

If defined, $[[\ sorry_i\ ]]^{f,g}(p)(a)(w)$=True iff $\forall w' \in max_{gi(\alpha,w)} f_i(\alpha,w): w' \notin p$

Now we can also give von Fintel's semantics for conditionals as well:

Admissible Modal Horizons
A function D from worlds to sets of worlds is an admissible modal horizon with respect to the ordering source g iff

$$\text{for any world } w,\ \forall w' \in D(w): \forall w''(w'' \leq_{g(w)} w' \rightarrow w'' \in D(w))$$

$[[\ would_i\ ]]^{D,g}$ (if p)(q)(w) is defined only if

(i) D_i is admissible with respect to g_i

(ii) $D_i(w) \cap p \neq \emptyset$

If defined, $[[\ would_i\ ]]^{D,g}$ (if p)(q)(w)=True iff $\forall w' \in D_i(w) \cap p$: q(w) = True.

Actually, it is not difficult to convince yourself that these meanings are Strawson AA. First, notice that being (Strawson) DE entails one direction of the equivalence that defines (Strawson) AA.

(101) For any A, B

(i) Suppose F is DE

(ii) $A \Rightarrow A \vee B$ theorem

(iii) $B \Rightarrow A \vee B$ theorem

(iv) So, $F(A \vee B) \Rightarrow F(A)$ by (i) and (ii)

(v) and $F(A \vee B) \Rightarrow F(B)$ by (i) and (iii)

(vi) Therefore, $F(A \vee B) \Rightarrow F(A) \wedge F(B)$

Therefore, to prove a DE function is AA all we have to prove is the other direction, i.e., $F(A) \wedge F(B) \Rightarrow F(A \vee B)$. Von Fintel has shown at length that adversatives and *would* conditionals are Strawson DE. All we need to show is that wide-scope conjunction Strawson-entails narrow-scope disjunction. This is easier to do than one might at first imagine. All we really need to do is look at the truth conditions. Why? Consider how Strawsonian reasoning works. To evaluate Strawson entailment you set up an argument whose premises include the truth conditions of the would-be entailer. If the truth conditions entail the truth conditions of the would-be entailee, the argument is valid. Why? It is a well-established result of logic that entailment is preserved under the addition of premises. Strawsonianism just adds the presuppositions of the arguments as premises.

The truth conditions of *sorry* say that the embedded proposition is disjoint from the set of g-best worlds in f. If P is disjoint from this set and Q is disjoint from this set, then $P \cup Q$ will be disjoint as well. So, $[[\ sorry_i\ ]]^{f,g}$(_)(a)(w) is a Strawson AA function.

The truth conditions for *would* conditionals say that the intersection of the antecedent proposition P with the modal base D is a subset of the consequent proposition R. If $P \cap D \subseteq R$ and $Q \cap D \subseteq R$, then $(P \cup Q) \cap D \subseteq R$ (note that $(P \cup Q) \cap D = (P \cap D) \cup (Q \cap D)$). So, $[[\ would_i\ ]]^{D,g}$ (if _)(q)(w) is a Strawson AA function.

APPENDIX 2: AA $\subseteq$ DE+Intolerant

Assume f is AA.

Suppose f is not trivial, i.e., $\exists x\ f(x)=1\ \&\ \exists x\ f(x)=0$.

Now suppose for reductio that $f(a)=1\ \&\ f(\neg a)=1$ for arbitrary a.

Notice that $a\vee\neg a = U$, that is, the top element in the domain.

Since f is AA it follows that $f(a\vee\neg a) = f(U) = 1$

But, being AA, f is DE. So, for all y such that $y\Rightarrow U$, $f(y)=1$.

But all y are such that $y\Rightarrow U$, so for all y, $f(y)=1$

(This contradicts our assumption that f is not trivial)

So, for all z, $f(z)=0$ or $f(\neg z)=0$.

Therefore, f is Intolerant.

References

Alonso Ovalle, Luis and Elena Guerzoni. 2004. Double negatives, negative concord and metalinguistic negation. In *The Proceedings of CLS 38.1: The Main Session*. Mary Androni, Erin Debenport, Anne Pycha and Keiko Yoshimura (eds.) Chicago Linguistic Society Publications, Chicago, IL. 15-31.

Atlas, Jay David. 1996. 'Only' noun phrases, pseudo-negative generalized quantifiers, negative polarity items, and monotonicity. *Journal of Semantics* 13:265–328.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In *Structures and Beyond*, ed. Adriana Belletti. Oxford: Oxford University Press.

Deprez, Vivienne. 1999 The roots of negative concord in French and French based creoles. In *Language Creation and Language Change: Creole, Diachrony and Development*, ed. Michel DeGraff. Cambridge, MA: MIT Press.

von Fintel, Kai. 1999. NPI-licensing, Strawson-entailment, and context-dependency. *Journal of Semantics* 16:97–148.

von Fintel, Kai. 2004. Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. In *Descriptions and Beyond*, edited by Marga Reimer & Anne Bezuidenhout. Oxford University Press.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.
Fox, Danny. 2003. The interpretation of scalar terms: Semantics or pragmatics, or both? Paper presented at the University of Texas, Austin.
Fox, Danny. 2007. Free Choice Disjunction and the Theory of Scalar Implicatures. In Uli Sauerland and Penka Stateva, editors, Presupposition and Implicature in Compositional Semantics, pages 71–120. Palgrave Macmillan, New York, NY.
Fox, Danny. 2008. Too many alternatives: density, symmetry, and other predicaments. In *Semantics and Linguistic Theory 17*.
Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.
Gajewski, Jon. 2005. *Only John* vs. *nobody but John*: A solution. Poster at NYU Polarity Workshop, March 11-12, 2005.
Gajewski, Jon. 2009. Innocent exclusion is not contradiction free. Ms. University of Connecticut.
Giannakidou, Anastasia. 2002. UNTIL, aspect and negation: a novel argument for two *untils.* in Brendan Jackson (ed.), *Semantics and Linguistic Theory 12*. CLC Publications, Cornell University, Ithaca, NY. 84-103.
Giannakidou, Anastasia. 2006. *Only,* emotive factives, and the dual nature of polarity dependency. *Language* 82:575–603.
Guerzoni, Elena. 2006. Intervention effects on NPIs and feature movement: Towards a unified account of intervention. *Natural Language Semantics* 14(4): 359-398.
Heim, Irene. 1984. A note on negative polarity and downward entailingness. In *Proceedings of NELS 14,* 98–107.
Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221
Hoeksema, Jacob. 1996. In days, weeks, months, years, ages: A class of negative polarity items. Ms.
Homer, Vincent. 2009. Disruption of NPI licensing: the case of presuppositions. In *Semantics and Linguistic Theory 18*. CLC Publications, Cornell University, Ithaca, NY. 429-446.
Horn, Laurence R. 1989. *A Natural History of Negation*. University of Chicago Press.
Horn, Laurence R. 1996. Exclusive company: *Only* and the dynamics of vertical inference. *Journal of Semantics* 13:1–40.

Horn, Laurence R. 2002. Assertoric inertia and NPI licensing. *Chicago Linguistics Society 38: Parassesion on Negation and Polarity*, 55–82.
Horn, Laurence R. To appear. Only connect: how to unpack an exclusive proposition. In a festschrift for Jay Atlas. Oxford University Press.
Karttunen, Lauri, and Stanley Peters. 1979. Conventional implicature. In *Syntax and Semantics 11: Presuppositions*, ed. C. Oh and D. Dineen, 1–56. New York: Academic Press.
Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25:1–49.
Krifka, Manfred. 1999. At least some determiners aren't determiners. In *The Semantics/Pragmatics Interface from Different Points of View*, ed. K. Turner. Elsevier Science B.V. 257–291.
Ladusaw, William A. 1979. *Polarity Sensitivity as Inherent Scope Relations*. Doctoral Dissertation, University of Texas, Austin.
Levinson, Dmitry. 2008. Licensing of Negative Polarity Particles *yet*, *anymore*, *either* and *neither*: Combining Downward Monotonicity and Assertivity. Doctoral Dissertation, Stanford University.
Linebarger, Marcia. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10: 325-387.
Löbner, Sebastian. 1985. Quantification as a major module of grammar. In *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, ed. J. Groenendijk, M. Stokhof, D. de Jongh. Dordrecht: Foris, S. 53–85.
Moltmann, Friederike. 1995. Exception sentences and polyadic quantification. *Linguistics and Philosophy* 18:223–280.
Nathan, Lance. 1999. Either: Negative polarity meets focus sensitivity. Ms., Brown.
Partee, Barbara H. 1989. Many Quantifiers. In *ESCOL 89: Proceedings of the Eastern States Conference on Linguistics*, eds. Joyce Powers and Kennith de Jong, 383-402. Department of Linguistics, Ohio State University: Columbus, OH.

Postal, Paul. 2004. The structure of one type of American English vulgar minimizer. In. *Skeptical Linguistic Essays*, Paul Postal. Oxford University Press.
Rooth, Mats. 1985. *Association with Focus*. Doctoral Dissertation, University of Massachusetts, Amherst.

Rullmann, Hotze. 2003. Additive particles and polarity. *Journal of Semantics* 20:329–401.
Sharvit, Yael and Simona Herdan. 2006. Definite and non-definite superlatives and NPI licensing. *Syntax* 9:1-31.
de Swart, Henriëtte. 1996. Meaning and use of *not…until*. *Journal of Semantics* 13:221–263.
van der Wouden, Ton. 1997. *Negative Contexts: Collocation, Polarity and Multiple Negation.* Routledge.
Zwarts, Frans. 1998. Three types of polarity. In *Plurality and Quantification*, eds. F. Hamm and E. W. Hinrichs. Kluwer: 177–238.