NPI *any* and Connected Exceptive Phrases

Jon Gajewski, University of Connecticut

Department of Linguistics, Unit 1145

337 Mansfield Road

Storrs, CT 06269-1145

Phone: (860) 681-3457

E-mail: jon.gajewski@uconn.edu

Abstract:

This paper addresses two puzzles in the semantics of connected exceptive phrases (EP): (i) the compatibility of EPs modifying noun phrases head by the negative polarity item (NPI) determiner *any* and (ii) the ability of a negative universal quantifier modified by an EP to license strong NPIs. Previous analyses of EPs are shown to fail to solve these puzzles. A new unified solution to the two puzzles is proposed. The crucial insight of the analysis is to allow von Fintel's (1993) leastness condition on EPs to be imposed at scopes non-local to the surface position of the EP. A general result is derived concerning the truth conditions of sentences in which an existential modified by an EP occurs in the scope of a downward entailing operator. The distribution of EPs is argued to depend on how the leastness condition interacts with other pragmatic strengthening conditions, such as the one imposed on *any* by Kadmon & Landman (1993).

1. Introduction

It is commonly held that a connected exceptive phrase (EP)[1] and its associated quantifier, such as *but John* and *no one* in (1), form a constituent at LF that denotes a generalized quantifier. Von Fintel 1993 proposes that EPs are Det-modifiers (i.e., of type <<et,ett>,<et,ett>>); Moltmann 1995 that they are DP-modifiers (i.e., of type <ett, ett>); Keenan & Stavi 1986 that Det + EP is a discontinuous determiner. In this paper, I present arguments against such "constituent" analyses. Instead I propose that (i) EPs are base-generated as sisters to NP (ii) EPs subtract the members of the set denoted by their complement from the set denoted by the NP and (iii) EPs are lexically marked to under go a specific kind of strengthening in the recursive pragmatics (cf. Chierchia 2004). I show that this approach to the semantics and pragmatics of EPs yields pleasing solutions to two difficult puzzles.

(1) $\underbrace{\text{No one}}_{\text{Associate}}\ \underbrace{\text{but John}}_{\text{EP}}$ ate the herring.

1.1 <u>Puzzle 1</u>: NPI *any* is compatible with EPs

There have been two major goals in the analysis of exceptive constructions: (i) to provide a unified account of their truth conditions and (ii) to explain the restricted distribution of EPs, preferably in terms of their semantics. The difficulty in achieving the first goal is explaining why (2a) implies that Bill is a student who smokes, but (2b) implies the he is a student who does <u>not</u> smoke.

(2) a. No student but Bill smokes.

b. Every student but Bill smokes.

The challenge behind goal (ii) is that EPs are very selective about the company they keep, associating only with universal quantifiers (with one notable exception).

(3) All/None/*Many/*Three/*Some/*Most of the students but Bill came

The theory that comes closest to meeting both goals is von Fintel (1993). Von Fintel formulates a uniform semantics for exceptive constructions that gives the correct truth conditions compositionally for EPs with both positive and negative universal associate quantifiers. Furthermore, his account predicts trivial truth conditions for (nearly all) sentences in which an EP associates with a non-universal quantifier. He suggests that the trivial truth conditions he predicts for the starred examples in (3) are sufficient to account for their ungrammaticality.

Von Fintel's account, however, predicts that trivial truth conditions, and therefore ungrammaticality, result <u>whenever</u> an EP associates with a non-universal quantifier. This is generally a welcome prediction (see (3)); but it is problematic in the case of the one exception to the universal generalization: NPI *any*.

NPI *any* and EPs do co-occur felicitously, as we in see (4).

(4) a. Mary didn't see anyone but Bill

b. No man saw any woman but Mary

In the literature, there are two main positions on the semantics of NPI *any*. Under one theory, that of Quine 1960 and Lasnik 1972 a.o., NPI *any* denotes a universal determiner that must take wide scope over negation. This would explain *any*'s compatibility with EPs straightforwardly. Under the other theory, that of Ladusaw 1979, Linebarger 1980 and Carlson 1981 a.o., NPI *any* denotes an existential determiner that must occur in the scope of a downward entailing (DE) operator. The thesis that NPI *any* denotes an existential determiner is by now well supported (see Kadmon & Landman 1993 a.o.). Advocates of this position have given a number of arguments that show NPI *any* must receive a narrow scope existential reading in some contexts. One such context is in the associate position of a *there* insertion sentence.

(5) There aren't any students in the room.

(6) a. *There aren't all the students in the room

b. *There isn't every student in the room.

(7) There must be someone in John's house. (Heim 1987)

(5) illustrates the acceptability of a DP headed by *any* as the associate of *there*. (6) shows that universal quantifiers are generally unacceptable as *there*-associates. (7) makes the further point that *there*-associates are generally restricted to narrowest scope: (7) does not have the reading that there is some particular person who is required to be in John's house. If *any* were always required to take wide scope over negation we would expect

(5) to be ungrammatical since *any* would be limited to narrowest scope.

Another argument that *any* is existential is based on the readings of sentences in which NPI *any* occurs in the scope of non-Anti-Additive DE quantifiers like *few*.

(8) An operator f is Anti-Additive (AA) iff $f(A \vee B) = f(A) \wedge f(B)$

(9) *few* NP is not AA since it is possible for (i) to be true and (ii) false.

(i) *few students smoke and few students drink*

(ii) *few students smoke or drink*

An AA operator makes narrow-scope disjunction (existentials) equivalent to wide-scope conjunction (universals). A non-AA DE operator does not. To see that *any* cannot be analyzed as a wide scope universal, it suffices to notice that (10a) can be false in a scenario in which (10b) is true with a wide-scope understanding of the object.

(10) a. Few students read any book on the list.

b. Few students read every book on the list.

In a scenario in which each book on the list was read by few students (10b) is true, but (10a) is false if the cumulative number of students who read some book or other on the list exceeds the contextual standard for fewness. For example, if each book was read by exactly three people, but the number of people who read from the list was 25, (10) would be false.

Given these facts about NPI *any*, we can sharpen the puzzle of its compatibility of EPs. NPI *any* is compatible with an EP even in these environments that argue for the existential (and against a universal) analysis of *any*. For example, consider the combination of NPI *any* with an EP in a *there*-insertion sentences (11a) and in the scope of the non-AA NPI licenser *few* in (11b).

(11) a. There isn't anyone in the room but Bill. (cf. Horn 1972)

b. Few boys talked to any girl but Sue.

PUZZLE 1: EPs associate felicitously with the existential NPI *any*

1.2 <u>Puzzle 2</u>: NPI licensing in the scope of [*no* NP *but* DP]

Turning to our next puzzle, notice that under any constituent analysis, the quantifier denoted by DP + EP will be non-monotonic, see (12). To be licensed, NPIs must occur in the scope of a downward entailing (DE) operator. Thus, such analyses predict that NPIs should not be licensed in the scope of quantifiers modified by EPs. This prediction is not borne out. For example, a negative universal does license NPIs in its scope when it is associated with an EP (13).

(12) a. $[\![$no student but Mary$]\!] = \lambda P.P \cap \{x: x \text{ is a student}\} = \{Mary\}$

b. $[\![$ every student but Mary $]\!] = \lambda P.\{x: x \text{ is a student}\} - P = \{Mary\}$

(13) No student but Mary said anything.

It is well known that non-monotonic quantifiers can license NPIs, at least in some contexts, see (14a). This kind of licensing, though, is heavily context-dependent, consider (14b).

(14) a. Exactly three students said anything in my seminar.

b. #Exactly ten of my twelve students said anything.

Non-monotonicity, however, is unlikely to explain the grammaticality of (13). Notice that the quantifier (12b) is every bit as non-monotonic as (12a) and yet it does not license NPIs at all:

(15) *Every student but Mary said anything.

Advocates of constituent analyses of EPs might take another tack. There are some apparently non-monotone quantifiers that do license NPIs systematically. One example is [*only* DP]. Semantically, [*only* DP] is quite similar to [*no* NP *but* DP] (and dissimilar to [*every* NP *but* DP]). Perhaps then, one could explain this licensing behavior by analogy to the NPI-licensing properties of the semantically analogous [*only* DP], see (17).

(16) a. Only Bill swam fast =/=> Only Bill swam

b. Only Bill swam =/=> Only Bill swam fast

(17) Only Mary said anything.

Von Fintel 1999 proposes, however, that [*only* DP] licenses NPIs because it is Strawson DE. While this provides an explanation for the licensing properties of [*only* DP], there is reason to not extend this analysis to [*no* NP *but* DP]. Interestingly, [*only* DP] and [*no* NP *but* DP] differ markedly in the classes of NPIs that they license. There is a class of NPIs that are not licensed by [*only* DP] (or non-monotonic quantifiers, or even merely DE operators for that matter) but are licensed by [*no* NP *but* DP]. These are the strong NPIs of Zwarts 1998. (on this contrast see Atlas 1993, Nathan 1999, Giannakidou 2006, a.o.)

(18) *Exactly three students left until five.

*Exactly three students have visited me in weeks.

Exactly three students like pancakes. *Exactly three students like waffles, either.

(19) *Only Bill left until five.

*Only Bill has visited me in weeks.

Only Bill likes pancakes. *Only Bill likes waffles, either.

(20) No student but Mary left until five.

No student but Mary has visited me in weeks.

No student but Mary likes pancakes. No student but Mary likes waffles, either.

The reputedly non-monotonic [no NP but DP] not only licenses NPIs, it licenses strong NPIs! This is a clear problem. A full discussion of these contrasts will be given in section 3.3.2.

> **PUZZLE 2**: [no NP but DP] is a strong NPI licenser.

1.3 Roadmap

In the remainder of this paper I will show that these two puzzles have a common solution. In section 2, I review previous approaches to the semantics of exceptives, finding that each fails to solve our two problems. In section 3, I show how a simple amendment to von Fintel's 1993 account can solve both puzzles. Section 4 concludes.

2. Previous Approaches to the Semantics of Exceptives

In this section, I review three analyses of EPs: von Fintel (1993), Moltmann (1995) and Postal (2000). I take special care in introducing von Fintel (1993), even though it offers no account of EPs with NPI *any*, since the proposal made in this paper is an elaboration of von Fintel's. Moltmann (1995) and Postal (2000) both offer specific accounts of NPI *any*'s compatibility with EPs. I offer criticisms of both. Then, in section 3, the alternative based on von Fintel (1993) is introduced.

2.1 von Fintel (1993)

Von Fintel (1993) gives an appealing account of the semantics of EPs. It offers a unified account of the truth conditions of EPs with positive and negative universals and a plausible explanation of the restriction on the distribution of EPs. On the other hand, it incorrectly predicts that EPs should not co-occur with NPI *any* and that EPs should not license NPIs.

2.1.1 Truth conditions

The basic truthconditional facts that von Fintel (1993) accounts for are the following:

(21) *every student but John complained* is true iff

(a) John is a student, and

(b) John did not complain, and

(c) every student who is not John complained

(22) *no student but John complained* is true iff

(a) John is a student, and

(b) John complained, and

(c) no student who is not John complained

Von Fintel treats each of the implications listed in (21) and (22) as a truthconditional entailment, since none of them is cancelable in the manner of an implicature or a presupposition.

Before looking at the details of von Fintel's analysis, it will be useful to introduce some terminology. Let's refer to a determiner D together with its restrictor A and scope P as a *quantification*. Given a quantification Q (=D(A)(P)), let's call a set C an *exception set* *relative to Q* if P belongs to D(A-C).

The first attempt that von Fintel makes at spelling out the truth conditions for EP constructions is (23). These truth conditions encode the basic idea that the complement of *but* denotes an exception set relative to the quantification in which the exceptive occurs, as defined above. An example is given in (24).

(23) [D A [[but]] C] P = True ⇔ P ∈ D(A–C)

↑ ↑ ↑ ↑ ↑

(24) *no boy but John ran* is True iff {x:x ran} ∈ [[no]]({x: x is a boy}– {John})

As von Fintel points out, (23) does not assign EP constructions strong enough truth conditions. Specifically, it does not account for the implications (a-b) in (21) and (22). Because of this weakness in truth conditions, (23) predicts counterintuitive entailments such as (25):

(25) no student but John complained ?⇒? no student but John and Mary complained

For this reason, von Fintel rejects (23) as inadequate. He proposes to keep domain subtraction as part of the meaning but to add another clause that will account for the additional implications. What he adds is the statement that the complement of *but* denotes the *least* exception set relative to its associated quantification. That is, the set denoted by the complement of *but* is a subset of every exception set of the quantification. This is formalized in the following way:

$$(26)\ [\text{D A}\ [\![\text{but}]\!]\ \text{C}]\ \text{P} = \text{True} \Leftrightarrow \underbrace{\text{P} \in \text{D(A–C)}}_{\text{Domain Subtraction}}\ \&\ \forall\underbrace{\text{S(P}\in\text{D(A–S)} \rightarrow \text{C}\subseteq\text{S)}}_{\text{Leastness}^{2}}$$

When this schema is applied to the determiners *every* and *no* we get (27) and (28) as special cases of (26). See von Fintel (1993) for more details on these proofs. (I use prime notation S′ to denote the complement of a set.)

$$(27)\ [\![\text{every}]\!]\ \text{A}\ [\![\text{but}]\!]\ \text{C P} = \text{True} \Leftrightarrow$$

$$\text{P}\in[\![\text{every}]\!](\text{A–C})\ \&\ \forall\text{S}(\text{P}\in[\![\text{every}]\!](\text{A–S}) \rightarrow \text{C}\subseteq\text{S})$$

$$\Leftrightarrow \text{A}\cap\text{P}'\subseteq\text{C}\ \&\ \forall\text{S}(\text{A}\cap\text{P}'\subseteq\text{S} \rightarrow \text{C}\subseteq\text{S})$$

$$\Leftrightarrow \text{A}\cap\text{P}'\subseteq\text{C}\ \&\ \text{C}\subseteq\text{A}\cap\text{P}'$$

$$\Leftrightarrow \text{A}\cap\text{P}' = \text{C}$$

$$(28)\ [\![\text{no}]\!]\ \text{A}\ [\![\text{but}]\!]\ \text{C P} = \text{True} \Leftrightarrow \text{P}\in[\![\text{no}]\!](\text{A–C})\ \&\ \forall\text{S}(\text{P}\in[\![\text{no}]\!](\text{A–S}) \rightarrow \text{C}\subseteq\text{S})$$

$$\Leftrightarrow \text{P}\cap\text{A}\subseteq\text{C}\ \&\ \forall\text{S}(\text{P}\cap\text{A}\subseteq\text{S} \rightarrow \text{C}\subseteq\text{S})$$

$$\Leftrightarrow \text{P}\cap\text{A}\subseteq\text{C}\ \&\ \text{C}\subseteq\text{P}\cap\text{A}$$

$$\Leftrightarrow \text{A}\cap\text{P}=\text{C}$$

These truth conditions straightforwardly yield the implications listed in (21) and (22), and bar entailments like (25).

2.1.2 Distribution: Ungrammaticality via "immediate falsity"

Let us turn now to how the truth conditions in (26) can account for the co-occurrence restrictions on EPs. Von Fintel's idea is that the "immediate falsity" at which one arrives by combining an EP with the wrong kind of quantifier results in ungrammaticality. The truth conditions in (26) require there to be a smallest set such that its subtraction from the restrictor of D makes the quantification true. On the one hand, universal determiners like *every* and *no* guarantee the existence of such least exception sets. Left upward monotone (↑mon) determiners, on the other hand, never have minimal exception sets.[3] That is, whenever a ↑mon determiner is substituted for D in the schema of (26), the right hand side of the equivalence is false, no matter what A, C and P are. Given that D is ↑mon and that $P \in D(A-C)$, we know that for any $S \subseteq C$, e.g. the empty set, $P \in D(A-S)$ since $A-C \subseteq A-S$. For this entire class of quantifiers, then, it is predicted that combination with an EP yields trivial falsity. And from triviality in truth conditions we make the leap to ungrammaticality. This idea has a precedent in Barwise & Cooper's 1981 analysis of *there* existential constructions.

Aside from some difficulties with proportional quantifiers in unnatural models (see Moltmann 1995, Lappin 1996 for discussion), this analysis is successful in picking out the positive and negative universal quantifiers as a class – the class of quantifiers that have least exception sets. In the next section, I show how von Fintel implements this proposed semantics in terms of lexical entries.

2.1.3 Compositional Implementation

Let's see how von Fintel cashes out his theory compositionally. From examining the truth conditions in (26) one sees that the denotation of *but* must be quite complex.

According to (26), in order to determine the truth value of an exceptive construction it is necessary to consider quantifications by D over subsets of A other than A-C. Consequently, according to von Fintel, the denotation of *but* must have independent access to the determiner (D) and its restrictor (A). That is, it must take them as separate arguments. The only flexibility in composition comes in deciding in which order the EP takes these arguments. The two syntactically plausible Schönfinkelizations of the denotation of *but* are (29a) and (29b):

(26) $[\text{D A } [\![\text{but}]\!] \text{ C}]\ \text{P} = \text{True} \Leftrightarrow \text{P} \in \text{D(A–C)}\ \&\ \forall \text{S}(\text{P}\in\text{D(A–S)} \rightarrow \text{C}\subseteq\text{S})$

(29) a. $\lambda C_{et}.\lambda A_{et}.\lambda D_{<et,ett>}.\lambda P_{et}.P\in D(A-C)\&\forall S(P\in D(A-S) \rightarrow C\subseteq S)$

b. $\lambda C_{et}.\lambda D_{<et,ett>}.\lambda A_{et}.\lambda P_{et}.P\in D(A-C)\&\forall S(P\in D(A-S) \rightarrow C\subseteq S)$

Von Fintel considers (29b) more plausible, since it implies that EPs are determiner modifiers. He regards the determiner-modifier analysis as possibly preferable because of the existence of constructions where EPs (30a), or the quasi-exceptive *almost* (30b), appear to modify a determiner directly.[4]

(30) a. All but at most five students came.

b. Almost all the students came.

The important point for us is that under either of these Schönfinkelizations, at LF, the constituent [*no* A but C] denotes the quantifier in (31), cf. (28).

(31) $[\![\text{no A but C}]\!] = \lambda P.A \cap P = C$

2.1.4 Problems for von Fintel 1993

Puzzle 1

Recall the examples in (4) and (11) illustrating the compatibility of EPs with NPI *any*. These sentences present a difficult challenge to von Fintel's theory. If *any* is an existential determiner, then it is $\uparrow$mon. As discussed in 2.1.2, under von Fintel's semantics, all $\uparrow$mon determiners yield trivial truth conditions when combined with an EP. So, von Fintel predicts incorrect (in fact trivial) truth conditions for such cases and furthermore predicts them to be ungrammatical.

Puzzle 2

According to a standard account (Ladusaw 1979), in order to be licensed NPIs must occur in the scope of a downward entailing operator. As can be seen in (32), NPIs are licensed in the scope of a negative universal quantifier associated with an EP.

(32) a. No student but Mary ever said anything.

b. No professor but Bill lifted a finger to help.

As illustrated in section 2.1.3, von Fintel predicts that the DP+EP in these sentences denotes a non-monotone quantifier. Thus, he predicts that NPIs should not be licensed in (32).

2.2 Moltmann 1995[5]

Moltmann 1995 offers a different approach to exceptives that is intended to cover a broader range of data than von Fintel's account. For example, Moltmann offers an analysis of quantified EP complements, as in (33), and of sentences involving irreducibly binary exceptions, as in (34).

(33) Every student but at most three law students left.

(34) a. Every man danced with every woman except Bill with Mary

b. No woman danced with any man except Sue with Fred

What is important for us is that Moltmann makes use of her analysis of both of these cases to explain the association of EPs with NPI *any*. So, let's go step by step through Moltmann's ideas.[6] Suppose for now that the complement of an EP denotes a set of individuals. First, Moltmann makes EPs into modifiers of generalized quantifiers (GQs), i.e., EPs map GQs to GQs. Second, Moltmann essentially imposes a presupposition that the associate of an EP is either positive universal or negative universal and that the set denoted by the EP complement is a subset of the associate's smallest live on set.

(35) A quantifier Q lives on a set A iff for every set X, ($X \in Q$ iff $X \cap A \in Q$)

The smallest live-on set of a GQ is its restrictor (if the GQ is semantically composed of (conservative) determiner and restrictor, cf. Barwise & Cooper 1981). Moltmann's basic idea is that if a DP is positive universal, then ⟦DP but S⟧ is the set of intersections of sets in ⟦DP⟧ with the complement of ⟦S⟧. In other words, take a set that is a superset of the set of boys, remove Billy and you've got a member of ⟦every boy but Billy⟧. If the DP is negative universal, then ⟦DP but S⟧ is the set of unions of sets in ⟦DP⟧ with ⟦S⟧. Take a set that has no boys in it, add Billy, and you have a member of ⟦no boy but Billy⟧. The semantics then is essentially disjunctive – one clause for *every*, one for *no*. This is an unattractive feature of Moltmann's approach.

Next, to deal with quantified EP complements, Moltmann extends this semantics pointwise over the witness sets of the quantified complement.

(36) A set w is a witness set for a quantifier Q iff for the smallest live-on set A of Q, $w \subseteq A$ and $w \in Q$

(37) W(Q) = the set of witness sets of Q

So for example, (38) is true if and only if (39) is true for at least one S where ⟦S⟧ is a witness set of ⟦at most one law student⟧. The witness sets of ⟦at most one law student⟧ are sets of law students with cardinality one or less. If there are two law students, Molly and Simon, then the witness sets are $\varnothing$, {Molly} and {Simon}. In that case, (38) is equivalent to (40).

(38) Every student but at most one law student left

(39) Every student but S left

(40) Every student left or every student but Molly left or every student but Simon left.

This correctly accounts for truth conditions of sentences like (38). So for Moltmann, a sentence with a quantifier EP complement Q is equivalent to the disjunction of sentences with EPs whose complements are witness sets of Q.

Finally, to deal with cases of binary exceptions like (34), Moltmann assumes that monadic quantifiers can be combined into dyadic quantifiers and that dyadic quantifiers can be modified by dyadic EPs. Moltmann extends the presupposition of (negative) universality to dyadic quantifiers, so that only dyadic quantifiers equivalent to the universals in (41) can be modified by EPs.

(41) a. $<[\![every]\!](A), [\![every]\!](B)> = \{R: A \times B \subseteq R\}$

b. $<[\![no]\!](A), [\![any]\!](B)> = \{R: A \times B \cap R = \emptyset\}$

A non-quantified EP complement now denotes a relation. The semantics is the same as with monadic quantifiers; [<every A, every B> but <C, D>] denotes the set of intersections of relations in $<[\![every]\!](A), [\![every]\!](B)>$ with the complement of C×D. Take a relation that contains every pair in A×B, take out any pair in C×D, and you have a member of [[<every A,every B> but <C,D>]].

When the binary EP complements are quantified, Moltmann uses the same strategy as above: pointwise extension of the semantics for non-quantified binary EPs over the witness sets of the quantified binary EP-complement. So, (42) is true if and only if (43) is true for some witness set S of ⟦at most one priest⟧. Suppose the only priests in the context are Father Jack and Father Eric. Then the three witness sets are ∅, {Father Jack} and {Father Eric}. In such a context, (42) is equivalent to (44).

(42) No man danced with any woman except at most one priest with Mary.

(43) No man danced with any woman except S with Mary.

(44) No man danced with any woman or

No man danced with any woman except Father Jack with Mary or

No man danced with any woman except Father Eric with Mary

Ingeniously, Moltmann sees a solution to our Puzzle 1 in her semantics for dyadic exceptives. What if, she wonders, EPs never directly modify a DP headed by NPI *any*? Suppose instead that any instance of an EP with NPI *any* involves dyadic quantification.

(45) No student looked at any paper but the Daily.

(46) logical form of (45)

[[but the daily](<no student, any paper>)](looked-at)

The immediate problem this analysis faces is that an EP that modifies a dyadic quantifier must have a relation-denoting expression as its complement. In (45), *the Daily* only supplies a set. Furthermore, as Moltmann points out, not just any relation will do. To get the interpretation of (45) right, the first term of the dyadic EP-complement must be an existential quantifier over <u>students</u>:

(47) No student looked any paper except one student at the Daily.

(48) logical form of (47)

[[but <one student, the daily>](<no student, any paper>)](looked-at)

The question of how *one student* is reconstructed within the EP is a crucial one for maintaining the compositional treatment of this construction. Moltmann suggests adding to the theory a semantics for EPs specifically for taking monadic exceptions to dyadic quantifiers. There are difficulties in assessing this proposal given its formulation.[7]

2.2.1 Problems for Moltmann (1995)

Compositionality

As already noted, Moltmann's account depends on an as yet unspecified account of how *but the Daily* comes to be interpreted as *except one student at the Daily*[8] in (45). As discussed in footnote 7, it is unclear given what Moltmann says, that a compositional semantics can be given to account for such cases as (45) that is not completely tailored to capture them. Notice, for example that we do not want to allow *but the Daily* to mean *except one student at the Daily* in other sentences such as (49).

(49) Every student looked at every paper but the Daily.

(50) Every student looked at every paper except one student at the Daily.

These sentences are distinct in meaning. (49) entails that no student read the Daily; (50) is compatible with students reading the Daily.

NPI licensing

In a sentence such as (51), if the EP modifies the monadic quantifier *no woman* then the result is a non-monotone quantifier. So, NPI-licensing is not predicted.

(51) No woman but Mary looked at any easy papers

A possible approach to these data open to Moltmann would be to say that *but Mary* does not modify *no woman* but the binary quantifier <no woman, any paper>. If this were so, the contribution of the exceptive would not interfere with *no woman* licensing *any paper*. There are two problems with this, however. First, it is infected with the compositionality problem since, under this analysis, (52) must receive the interpretation:

(52) No woman looked at any easy papers but Mary <u>at one easy paper</u>.

Second, as Moltmann observes (1995, footnote 27) exceptives that apply to true dyadic quantifiers must be extraposed:

(53) Every man danced with every woman except Jim with Mary.

(54) *Every man except Jim with Mary dance with every woman.

Why extraposition is not necessary in (51) would require an explanation.

Indefinites other than NPI *any*

Moltmann's analysis of EPs with NPI *any* would seem to extend straightforwardly to other indefinites in the scope of negation. For example, (55) is acceptable. It is a surprise then that (56) is not acceptable.

(55) Last year in this class, no one read a novel on the reading list.

(56) *No one read a novel on the reading list but Moby Dick.

This is puzzling since <⟦no student⟧, ⟦any novel⟧> = <⟦no student⟧, ⟦a novel⟧> so long as ⟦ any novel ⟧ = ⟦ a novel ⟧. Moltmann suggests (1995, p.275) that the syntactic connection between a negation and its licenser may play a role in allowing them to form a dyadic quantifier. While this may be, NPI-licensing cannot be a general precondition on dyadic quantifier formation since there is no licensing involved in sentences like (53). So, why couldn't *a novel* form a dyadic quantifier with *no student*?

Similar comments apply to *some*. Though *some* cannot generally appear in the scope of negation, in those cases when it can it still cannot host an EP.

(57) Bill didn't read some novel on the reading list *NOT>SOME

(58) You can't tell me Bill didn't read some novel NOT>SOME

(cf. Baker 1970)

(59) *You can't tell me Bill didn't read some novel but Moby Dick

What accounts for the impossibility of forming the dyadic quantifier <⟦not⟧, ⟦some novel⟧> in (59)?

2.3 Postal (2000), (2005)

Though Postal (2000),(2005) does not explicitly discuss the semantics of exceptives, he does sketch a new account of the relation between *no* and *any* that may shed light on our puzzles. Postal's proposal is particularly interesting because it avoids the problem just noted for Moltmann (a challenge for my analysis also, as we will see). Postal avoids this problem by denying that *any* uniformly denotes an existential quantifier. Instead, according to Postal, NPI *any* may denote an existential quantifier or a negative universal quantifier, depending on the context. One piece of evidence he presents for this is the distribution of EPs. Postal wishes to maintain the strong generalization that EPs only associate with (negative) universal quantifiers. Thus he is driven to say that in the cases that NPI *any* associates with an EP it denotes a negative universal quantifier. Consequently, he must complicate the mapping to morphology so that an underlying negative universal is sometimes spelled out as *no* and sometimes as *any*.

Postal suggests [$_D$NEG SOME] as the underlying representation of the negative universal that may be spelled out either as *any* or *no*. NEG and SOME denote cross-categorial negation and an existential determiner, respectively. The existential that sometimes underlies *any* is represented as [$_{\text{D}}$NEG[$_{\text{D}}$NEG SOME]], semantically equivalent to *some*.

The spell out procedure proceeds as follows. First, the rule (60) applies to every instance of SOME that shares a D node with NEG at any point in the derivation.

(60) [$_{\text{D}}$NEG SOME] ⇒ [$_{\text{D}}$NEG ANY]

Second, a rule that applies at the surface:

(61) [$_{\text{D}}$NEG ANY] ⇒ [$_{\text{D}}$NO]

This rule must wait to apply at the surface because it can be bled by NEG-deletion. There are two rules of NEG-deletion relevant for our purposes:

(62) [$_{\text{D}}$NEG[$_{\text{D}}$NEG ANY]] ⇒ [$_{\text{D}}$NEG[$_{\text{D}}$~~NEG~~ ANY]]

(63) [$_{\text{D}}$NEG ANY] ⇒ [$_{\text{D}}$~~NEG~~ ANY], if NEG is c-commanded by a local AA operator

(64) [$_{\text{D}}$NEG[$_{\text{D}}$NEG ANY]] ⇒ [$_{\text{D}}$~~NEG~~[$_{\text{D}}$NEG ANY]], if NEG is c-commanded by a local DE operator

So, then, an underlying [$_D$NEG SOME] can host an exceptive because it is semantically a negative universal, yet it can be spelled out as *any* if it is licensed by an AA operator. Another [$_D$NEG SOME] could be such an AA operator. This, however, can't be right – the meaning comes out wrong.

(65) No student read any book but Moby Dick

(66) a. NEG1-SOME student read NEG2-SOME book (but Moby Dick)

b. NEG1-ANY student read NEG2-ANY book (but Moby Dick) (62)

c. NEG1-ANY student read ~~NEG2~~-ANY book (but Moby Dick) (63)

"no student read any book but Moby Dick" (spell-out)

This is clearly wrong, as Postal notes. If this were the correct underlying structure, the sentence would mean no student read <u>no</u> book but Moby Dick. To correct this, Postal postulates an unpronounced negation modifying the verb. The presence of such a negation is required to satisfy a global constraint that only an even number of NEGs can be deleted. This verb negation, being AA, can delete the NEG of the object's determiner. A rule like (67) must be added to license deletion of the NEG on the verb:

(67) NEG+V ⇒ ~~NEG~~+V, if NEG is c-commanded by an AA operator

(68) a. NEG1-SOME student NEG2-read NEG3-SOME book (but Moby Dick)

b. NEG1-ANY student NEG2-read NEG3-ANY book (but Moby Dick) (62)

c. NEG1-ANY student NEG2-read ~~NEG2~~-ANY book (but Moby Dick) (63)

d. NEG1-ANY student ~~NEG2~~-read ~~NEG2~~-ANY book (but Moby Dick) (67)

"no student read any book but Moby Dick" (spell-out)

2.3.1 Problems for Postal

Truth conditions

Unfortunately, (68) is also wrong. Postal predicts that (65) is equivalent to (69), which it is not.

(69) Every student read no book but Moby Dick.

The point of Postal's proposal is to be able to treat EPs with NPI *any* as simple EPs modifying a (negative) universal monadic quantifier. But all approaches to the semantics of monadic EPs make incorrect predictions about the truth conditions of (65), if it is assigned the underlying structure in (68). Specifically, it is predicted that (65) entails, as (69) actually does, that every student read Moby Dick. (65) does not entail this, it merely entails that <u>some student</u> read Moby Dick.

I do not see a simple way to derive a structure in Postal's system that is pronounced as in (65), but has truth conditions suitably weaker than (69). Perhaps, Postal could get some mileage out of the fact that (65) is equivalent to (70). The spell-out system, however, would have to be revised to accommodate QR.

(70) No book but Moby Dick has any student read.

NPI-licensing

Postal 2005 notes that an AA operator is needed to license the surface form *any* when it is expanded with *not even* X.

(71) No girl but Molly danced with any boy, not even Simon.

The well-formedness of (71) suggests the presence of an AA operator. However, if the EP directly modifies *no girl* as Postal implies, there will be no AA operator c-commanding *any boy not even Simon*, only the non-monotonic *no girl but Molly*.

2.4 Conclusion

In this section we have seen that no existing approach to EPs offers an adequate solution to either of our two Puzzles. Moltmann's (1995) proposal suffers from three problems: the lack of a general rule for forming dyadic EPs from surface monadic EPs, an incomplete explanation of NPI licensing by EP-modified GQs, overgeneration of EP sentences – predicting compatibility with all indefinites in the scope of *no*. Postal's (2000), (2005) proposal offers an interesting solution to Moltmann's indefinite problem but derives incorrect truth conditions for the crucial sentences, e.g. (65). Postal also predicts that EP-modified GQs have weaker licensing properties than they display. To my mind, the most promising of these accounts is von Fintel's, since he offers a unified semantics and an explanation of the distribution of EPs that follows naturally from the semantics he proposes. In the next section, I argue that an amendment to the compositional implementation of von Fintel's semantics offers a common solution to our two Puzzles.

3 Dividing up the meaning of Exceptives

The guiding light for my approach to EPs is von Fintel (1993) and the idea that its Leastness condition on EPs is rooted in a kind of pragmatic reasoning. This leads me to propose that the Leastness part of an EPs meaning be separated from Domain Subtraction and enforced by a strengthening operator of the recursive pragmatics (cf. Chierchia 2004).

In section 3.1, I present my revision of von Fintel's account. In essence, I retain the semantics for EPs proposed by von Fintel (1993). I propose to change, however, the compositional implementation of the semantics. This change splits the scope at which the two parts of von Fintel's semantics apply.[9] The Domain Subtraction is enforced at the surface position of the EP. The Leastness of the EP-complement, however, may be imposed at a point higher in the tree. The driving analogy behind the non-locality of Leastness is with Chierchia's (2004, 2006) pragmatic strengthening operators and in particular his account of the licensing of *any*. In section 3.2, I give my account of the truth conditions of sentences in which an EP co-occurs with NPI *any*. In section 3.3, I show how this analysis provides a natural account of the NPI licensing properties of [*no* A *but* C]. Finally, in section 3.4, we return to the issue of ruling out EPs with non-NPI indefinites in the scope of DE operators.

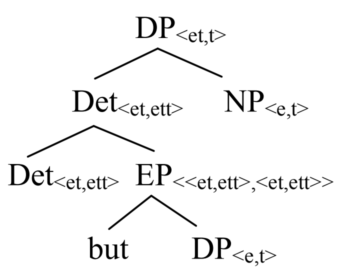
3.1 Non-Local Leastness: A Revision of von Fintel 1993

Recall the problem confronting von Fintel's analysis. Von Fintel assigns the meaning schema in (26) to sentences containing exceptive phrases. He implements this compositionally by assigning *but* the denotation in (29b), resulting in structures of the form (72).

(26) $[\text{D A } [\![\text{but}]\!] \text{ C}] \text{ P} = \text{True} \Leftrightarrow \text{P} \in \text{D(A–C)} \;\&\; \forall \text{S}[\text{P} \in \text{D(A–S)} \rightarrow \text{C} \subseteq \text{S}]$

(29b) $[\![\text{but}]\!] = \lambda \text{C}_{et}.\lambda \text{D}_{<et,ett>}.\lambda \text{A}_{et}.\lambda \text{P}_{et}.\ \text{P} \in \text{D(A–C)} \& \forall \text{S}[\text{P} \in \text{D(A–S)} \rightarrow \text{C} \subseteq \text{S}]$

(72)



Applying (26) to an ↑mon determiner results in immediate falsity. Thus, if NPI *any* denotes an existential determiner, then [*any* NP *but* DP] ought to be ungrammatical due to immediate falsity. Such a combination of NPI *any* and *but* is not ungrammatical, e.g., (73). Such sentences are grammatical and have clear entailments that a theory of EPs should account for:

(73) *John didn't see any student but Mary* is true iff

(i) Mary is a student

(ii) John saw Mary

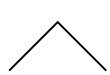
(iii) John didn't see a student who was not Mary

I suggest that von Fintel's analysis can capture these entailments, if we give it a different compositional structure. As was noted in section 2.1.3, von Fintel considers only two ways of assigning a meaning to *but*, (29a) and (29b). These are indeed the only two syntactically plausible possibilities under the assumption that Domain Subtraction and Leastness are imposed together with the same scope. If we abandon this assumption, we see that there are other ways to slice the meaning pie of (26) compositionally.

Notice first that *but* need not take D, A, C and P individually as arguments. The only one of these that is crucially singled out in the truth conditions is C, the exception set. The others, D, A and P, are referred to in both conjuncts as a unit, namely the function $\lambda X. P \in D(A-X)$: in the first conjunct, it is stated that $P \in D(A-C)$; in the second conjunct, '$P \in D(A-S)$' is used in the antecedent of a conditional, where S is a variable bound by a higher quantifier. I propose that we can obtain just these two pieces if we make the following assumptions: (i) *but* denotes set subtraction and (ii) *but* introduces alternatives to its complement. I use the mechanisms of focus-marking to introduce alternatives, and assume a structured meaning approach to focus, see Jacobs 1983, Krifka 1991. I am not making any claims about focus on the complement.

(74) $[\![\text{but}]\!] = \lambda B.\lambda A.A-B$

(75) *but* obligatorily marks its complement with an F.

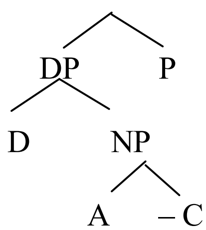


but DP_F

(76) $[[\ A_F\ ]] = \langle \lambda X.X, A \rangle$

Then, we obtain (78) as the focus denotation of (77).

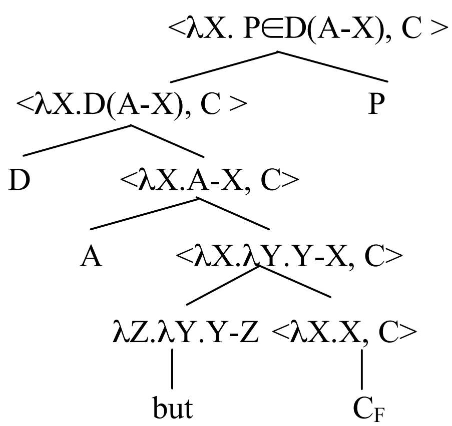
(77)



e.g., [[no student but Mary] left]]

D A C P

(78)



(79) For example the focus denotation of *no student but Mary$_F$ left* =

$\langle \lambda X.\ \{y{:}y \text{ left}\} \in [[\text{no}]](\{x{:}x \text{ is a student}\} - X), \{\text{Mary}\} \rangle$

Given these assumptions, I propose that an EP is required to be bound at LF by a kind of pragmatic strengthening operator. As von Fintel notes in laying out his semantics for exceptives,

> It should be obvious that the uniqueness [leastness – jrg] condition is pragmatically natural. It ensures maximal relevance of the *but*-phrase: the exceptive is not only necessary to save the quantification, it is also the most economical way of doing that.

Whereas von Fintel takes the Leastness condition to be built into the semantics of EPs, I assume it still has some freedom to apply in larger domains. Specifically, I propose that there is a clausal operator LEAST that takes a structured meaning as its argument, where the second element of the structured meaning is a set of individuals.

(80) [but DP_F] must be c-commanded at LF by LEAST

(81) $\text{LEAST}(\langle F,X\rangle) = 1 \text{ iff } F(X)=1 \;\&\; \forall S[\, F(S)=1 \rightarrow X\subseteq S \,]$

I propose that LEAST belongs to the same class of operators as Chierchia's (2004, 2006) **O** and σ, Fox's (2006) EXH and Kadmon's (1987) maximality operator.[10] Furthermore, I propose that assigning LEAST the appropriate scope will solve our two Puzzles.[11]

Consider now a structure in which LEAST immediately c-commands a simple quantificational structure made up out of a determiner, its restrictor and its scope, as in (77) above.

(82) LEAST [D [A *but* C_F] P]

(83) LEAST(<λX.D(A-X)(P)=1, C>) = 1 iff

$$D(A\text{-}C)(P) = 1 \ \& \ \forall S[\ D(A\text{-}S)(P)=1 \rightarrow C \subseteq S]$$

As desired, then, we have replicated von Fintel's results for simple structures like those in (77). The power of our new analysis, though, is that LEAST need not only appear in configurations of the form in (77). Typewise, it can apply to any node of type t dominating the EP. Nothing we have said so far dictates *which* node of type t LEAST applies to. It might take more structure into its scope than the quantifier in which its associated EP occurs.

3.2 PUZZLE 1: PART 1

This freedom in scope sites for LEAST gives us the flexibility we need to capture the truth conditions of examples with NPI *any*. Let's consider (73) as a concrete example. Under von Fintel's (1993) proposed denotation for *but*, the EP was constrained to take the determiner *any* as an argument, resulting in trivial truth conditions. We showed above that we make the same predictions as von Fintel (1993) when LEAST takes scope as in (82). So, we predict that an LF is ill-formed in which LEAST takes scope just above the existential NPI *any student* and below sentential negation.

(84) *LF of (73) = [not [LEAST [any [student – $Mary_F$] 2 [John saw t_2]]]]

Here the constituent that is sister to sentential negation is of the same form as (77) and therefore induces ungrammaticality because of trivial truth conditions. We do not yet predict, however, that (73) is ungrammatical simpliciter – other LFs are available. The other scope option that is available to LEAST is the root node. This is a scope option that is not of the form (77): one where there is another operator in the scope of [but Mary], namely negation. Here then we make different predictions from von Fintel 1993.

(85) Well-formed LF for (73): [LEAST [not [[any student – $Mary_F$] 2[John see t_2]]]]

The reason that this LF is well formed is intuitively clear. The denotation of the sister of LEAST is equivalent to the denotation of a constituent that *does* have the form (77): namely one in which D = *no*.

(86) $[\![$ not [[any student – Mary] 2[John saw t_2]] $]\!]$ = $[\![$ [no student – Mary] 2[John saw t_2] $]\!]$

Given this equivalence we can fall back on von Fintel's analysis to determine the truth conditions of (73). For the sentence with *no*, von Fintel makes the following predictions:

(87) *John saw no student but Mary* is True iff

$\{x: x \text{ is a student}\} \cap \{x: \text{John saw } x\} = \{\text{Mary}\}$

These are exactly the truth conditions we want for (73), since they account for the entailments in (73i-iii).

3.2.1 Beyond *not…any*

This analysis is equally successful with cases that cannot be so easily reduced to structures for which von Fintel (1993) makes correct predictions. One such case is (88).

(88) *No student read any book but War & Peace* is true iff

(i) W&P is a book, and

(ii) *Some* student read W&P, and

(iii) No student read a book that wasn't W&P

Similarly to (73), the LF of (88) in which LEAST scopes below the NPI licensor is ill-formed due to its trivial truth conditions. That is, (89) is filtered out at LF.

(89) *LF of (88): [no student 1[LEAST [any [book – WP_F] 3[t_1 read t_3]]]]

The other scope possibility for LEAST, above *no student*, does yield a well-formed LF:

(90) Well-formed LF for (88): [LEAST [no student 2[any [book – WP_F] 3[t_2 read t_3]]]]

The truth conditions we assign to this structure are given in (91). The import of these conditions may not be immediately obvious.

(91) $[\![(90)]\!] = 1$ iff

(i) $[\![\text{student}]\!] \cap \{x: \{y: x \text{ read } y\} \cap [\![\text{book}]\!] - \{WP\} \neq \emptyset\} = \emptyset$

(ii) $\forall S[[\ [\![\text{student}]\!] \cap \{x: \{y: x \text{ read } y\} \cap [\![\text{book}]\!] - \{WP\} \neq \emptyset\} = \emptyset] \rightarrow \{WP\} \subseteq S\ ]$

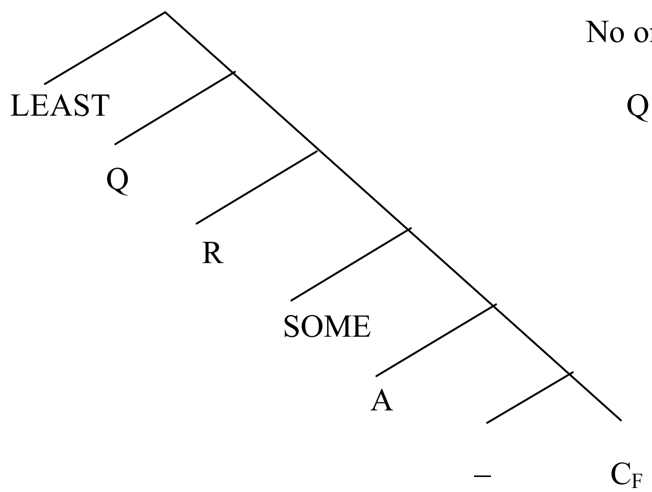
On our way to a more general and understandable result, let's note that the conditions in (91) are equivalent to those in (92).

(92) $[\![(90)]\!] = 1$ iff $\{WP\} = \{y: y \in [\![\text{book}]\!]\ \&\ [\![\text{student}]\!] \cap \{x: x \text{ read } y\} \neq \emptyset\}$

I now informally sketch the equivalence of (91) and (92). If (91i) is true, then if there are any books that were read by students, then they must be contained in {WP}, i.e., $\{y: y \in [\![\text{book}]\!]\ \&\ [\![\text{student}]\!] \cap \{x: x \text{ read } y\} \neq \emptyset\} \subseteq \{WP\}$. Now suppose WP was read by no students, i.e. {WP} is not a subset of the set of books read by students. Then, by (91ii) it must be false that no student read a book that no student read. But that is a contradiction, so we conclude that $\{WP\} \subseteq \{y: y \in [\![\text{book}]\!]\ \&\ [\![\text{student}]\!] \cap \{x: x \text{ read } y\} \neq \emptyset\}$. And therefore, $\{WP\} = \{y: y \in [\![\text{book}]\!]\ \&\ [\![\text{student}]\!] \cap \{x: x \text{ read } y\} \neq \emptyset\}$. Now in the other direction, if {WP} is the set of books read by students, then (91i) is clearly true – it is a tautology that no student read a book that no student read. (91ii) follows, as well. If the set of books read by a student is not a subset of another set S, that must be because there is a book read by a student that S does not contain. But then when S is subtracted from the set of books, that student-read book stays in the domain making the antecedent of the conditional in (ii) false. So, the equivalence is proven.

In the same vein, there is a more general result that holds for structures such as (93), where Q is Downward Entailing and SOME stands in for any existential quantifier:

(93)



No one read any book but WP_F

Q R SOME A – C

(94) $[\![(93)]\!] = 1$ iff $\{x: \{y: xRy\} \in SOME(A–C)\} \in Q$ &

$C = \{z: z \in A \ \& \ \{y: yRz\} \notin Q\}$

A formal proof of this result is given in the Appendix. [As expected, if Q in (93) is upward entailing, the sentence is ruled out for having trivial truth conditions, see Appendix for proof.] This general result will make it easier to understand the truth conditions we get for other complex cases.

One last word about (90): its truth conditions appear simpler than those in (94). The reason for this is that, in this case, the first conjunct of the truth conditions of (94) is entailed by the second (this is not always the case, see below). The second conjunct alone says that War & Peace is the one and only book that some student read. This is exactly what we want as it is equivalent to the conjunction of the truth conditions listed in (88).

Let's now use our general result to calculate the truth conditions for an example of the form (93) in which Q is DE but not AA. Take (95).

(95) Few students read any novel but Ethan Frome.

(96) a. *Few students$_1$ LEAST [[any novel – EF]$_2$ t_1 read t_2]

b. LEAST [Few students$_1$ [[any novel – EF]$_2$ t_1 read t_2]]

(97) $[\![\text{(96b)}]\!] = 1$ iff $\{x: \{y: x \text{ read } y\} \in [\![\text{some}]\!]([\![\text{book}]\!] - \{WP\})\} \in [\![\text{few student}]\!]$ &

$\{EF\} = \{z: z \text{ is a book } \& \ \{x: x \text{ read } z\} \notin [\![\text{few student}]\!]\}$

Here, the second conjunct of the truth conditions does not entail the first. It could be true that Ethan Frome is the only novel that many (i.e. not few) students read and false that few people are such that they read a novel that is not Ethan Frome. So long as every other individual novel was read by few people, as many people as you like can have read one of them and the second conjunct of (97) would still be true. The first conjunct keeps the number of readers of non-Ethan Frome books down to few. These predictions seem empirically correct, as the following two scenarios show.

First, here is a scenario in which we predict (95) to be true (suppose less than 3 counts as few):

(98) There is a class with 10 students. The reading list for the class has four books on it: Ethan Frome EF, The Age of Innocence AI, House of Mirth HM & Glimpses of the Moon GM. Bill read the AI and EF, Sue read HM and EF and Sara read EF. No other students read any other books.

Next, here is a scenario in which we predict (95) to be false:

(99) Same class and reading list as in (97). Fred read AI. Sue read AI. Bill read HM. Sara read HM. Ted read GM. Mary read GM. Everyone read EF.

Indeed, (95) is false in this scenario since nearly everyone read a book that wasn't Ethan Frome. This scenario is of interest since we would predict that (95) should be true in (99) if we analyzed it with a wide scope universal, i.e., as being equivalent to (100):

(100) Every book but Ethan Frome is such that few students read it.

Conclusion

Consider what we have accomplished here. By making a small amendment to von Fintel's semantics for EPs – by allowing non-local enforcement of Leastness – we have been able to give a (perhaps the first) fully compositional account of the truth conditions of EPs with NPI *any*. There is no need to construct a dyadic exception out of monadic cloth or assume a complex mapping from semantics to morphology.

Consider what we have not yet accomplished. Much like Moltmann's, our proposal, as it stands, predicts that sentences like (56) in which *a* replaces *any* ought to be every bit as acceptable as (88) itself. As we have seen this is not the case. More generally, something must be said about what the possible scope sites are for our LEAST operator. I address this issue in section 3.4.

3.2.2 LEAST vs. ONLY

The reader may wonder why we must resort to the highly grammaticalized LEAST operator when an oft-invoked, close pragmatic cousin lies at hand, namely the ONLY-based strengtheners of Fox 2006 and Chierchia 2006.

(101) $\text{ONLY}\langle F,D\rangle = F(D) \text{ and } \forall S[F(S) \rightarrow F(D) \Rightarrow F(S)]$

Empirically, ONLY cannot be the strengthener involved with EPs, since it does not yield contradictions for $\uparrow$mon determiners.

(102) $\text{ONLY}[\ [\ \text{SOME A} - C_F\ ]\ P\ ]$

$\text{ONLY}\langle \lambda S.\text{SOME}(A\text{-}S)(P), C\rangle$

Since *some* is $\uparrow$mon, if the statement is true when you take a larger exception, it will be true when you take a smaller exception. So, the use of ONLY with *some* would require the complement of *but*, C in (102), to denote the greatest exception set. There will be a greatest exception set when the intersection of the restrictor A and the scope P is a singleton; the greatest exception set is then the complement of that singleton in the restrictor, A–(A∩P), i.e., A–P.[12]

To take a concrete example, if *but* could be strengthened with ONLY, we would predict that (103) could be true and, therefore, is grammatical. The reason is that (103) would be true in a world where (104) holds. In such a world, {Sue, Fred} is the greatest exception set.[13]

(103) *Some student but Sue and Fred smokes

(104) a. [[student]] = {Phil, Sue, Fred}

b. [[smokes]] = {Phil}

As a reviewer observes, this may be the correct analysis for the closely related *other than* construction, which is compatible with *some*. (105) can, but need not, imply that Bill did not arrive.

(105) Some student other than Bill arrived.

Another difference between ONLY-like operators as discussed by Fox and Chierchia and our LEAST, is freedom in scope-taking. Fox and Chierchia allow their strengthening operators to take non-minimal scope. We can actually observe the difference between LEAST and ONLY in a near minimal pair. *But* contrasts minimally here with *other than*.

Other than statements often imply that their complements are true exceptions, but this is defeasible. As shown by the fact that (107) is a felicitous continuation of (106).

(106) Every student other than Phil finished the exam.

Implies "Phil didn't finish the exam"

(107) Maybe Phil did, too. I haven't checked yet.

Now notice that when *other than* is embedded under another operator Op the strengthening can be local (within the scope of Op) or non-local (taking scope over Op).

(108) Mary is certain that every student other than Phil finished the exam.

a. Mary is certain that ONLY[every student other than Phil finished the exam]

b. ONLY[Mary is certain that every student other than Phil finished the exam]

This sentence can implicate the strong conclusion that Mary is certain that Phil didn't finish the exam, but it need not. It can also implicate the weaker conclusion that Mary is not certain that Phil finished. This contrasts crucially with *but*: (109) entails that Mary is certain that Phil has not finished.

(109) Mary is certain that every student but Phil has finished the exam.

This leaves the question of why *but* requires strengthening by LEAST, rather than ONLY. Here I can only speculate. Notice that LEAST entails that the statement containing the exceptive is false. For example, *no student but Mary left* entails that *no student* left is false. This is generally true of LEAST, whenever the exception set is non-empty it is entailed that subtracting the empty set would yield falsity. ONLY does not have this property as demonstrated by its yielding consistent truth conditions for *some*. The truth conditions of that output are compatible with the truth conditions of the unmodified statement *some student arrived*. The reason that *but* requires strengthening by LEAST, rather than ONLY, may be that this accords better with an EPs function as a repair for false quantifications.

3.2.3 Too far beyond?

As noted above, we predict that a sentence of the form (110) has the truth conditions in (111), where Q is a DE generalized quantifier and R a transitive verb.

(110) LEAST[Q 1[SOME(A-C_F)2[t_t R t_2]]]

(111) $[\![$ (110) $]\!]$ = 1 iff {x: {y: xRy}∈SOME(A–C)}∈ Q &

C = {z: z∈A & {y: yRz}∉Q}

For most choices of a DE Q, it is a simple matter to create a model that shows that these truth conditions are consistent. For example, we predict consistent truth conditions and thus grammaticality for each of the sentences in (112).[14] Specifically, we predict that (112a) is true if and only if not every student read a book that is not Ethan Frome and every student read Ethan Frome.

(112) a. ?[Not every student] read any book but Ethan Frome

b. ?[Less than three students] read any book but Ethan Frome

c. ?[At most two students] read any book but Ethan Frome

This is a prima facie surprising result. Von Fintel's (1993) analysis very nearly picks out only the left anti-additive determiners as yielding non-trivial truth conditions with EPs (in simple constructions of form (77)) .[15] Under our extension, however, we see that von Fintel's semantics can yield consistent truth conditions for merely DE functions like [λX. few students read any X], denoted by more complex constructions of the form (93). That is, while (113a) is predicted – by von Fintel and me – to be ungrammatical, (113b) is not. (114) shows that the function denoted by the environment of the EP in (113b) is not AA.

(113) a. *Few students but Bill arrived early

b. Few students read anything but War & Peace.

(114) Few students read any journals and Few students read any novels =/⇒

Few students read any [journals or novels]

This suggests there is a constraint on the semantics of determiners that rules out the kind of merely DE functions that yield consistent truth conditions under von Fintel's semantics. Again see, Moltmann (1995) and Lappin (1996) for discussion.

I even predict that it possible for an exceptive to yield consistent truth conditions when associated with NPI *any* in the scope of a non-monotonic quantifier, to the extent that NPI-licensing is possible. Consider a scenario, in which there are two relevant books: Ethan Frome which was read by exactly four students and The House of Mirth which was read by exactly three. Then, it is necessary and sufficient to remove Ethan Frome from the domain of books to make the statement *Exactly three students read any book* true.

(115) ??Exactly three students read any book but Ethan Frome.

If it is deemed that we should exclude (112) and (115), perhaps the avenue to explore is the difference between allowing least exception sets and guaranteeing their existence (see discussion in von Fintel 1993, pp.132-133). The realm of quantificational determiners is (nearly) divided between those that guarantee least exception sets and those that uniformly exclude them. In the wider worlds, there are many functions that fall in between. (112) and (115) exhibit cases that allow, but don't guarantee, least exception sets. (88), on the other hand, guarantees that a least exception set exists. The account of which cases are ungrammatical would have to become more indirect. I will continue assume the explanation of ungrammaticality of EPs in terms of triviality. In the section, I briefly explicate the difference between allowing and guaranteeing least exception sets for readers that find the sentences in (112) unacceptable.

3.2.4 Guaranteeing Least Exception Sets

An EP is useful when the statement it modifies is false. If the statement is true there is no need for the exceptive. We have followed von Fintel (1993) in assuming that the complement of *but* must denote the least exception set. The least exception set for a statement of the form (116) is $\{x{:}x{\in}A \;\&\; {\sim}Q(\lambda y.yRx)\}$, given that we have calculated the truth conditions for (116) as (117) – when Q is DE.

(116) $\text{LEAST}[\ Q\ 1[\ \text{SOME}(A\text{-}C_F)2[\ t_t\ R\ t_2\ ]\ ]\ ]$

(117) $Q(\lambda x.\text{SOME}(A\text{-}C)(\lambda y.xRy))=1\ \&$

$$C = \{x{:}x{\in}A \;\&\; {\sim}Q(\lambda y.yRx)\}$$

For example if the original statement is (118), the least exception set (on the domain of the object quantifier) is (119).

(118) No student read any book.

(119) {x: x is a book & some student read x}

Notice now that the falsity of (118) implies that (119) is non-empty. If it is false that no student read any book, then there must be books that some student read.

When the Q in a statement of form (116) is DE but not AA, there is no guarantee that the least exception set is non-empty when the unmodified statement is false. Compare the case of (120) with (118).

(120) Not every student read any book.

(121) {x: x is a book and every student read x}

The least exception set is (121). The falsity of (120), however, does not guarantee that (121) is non-empty. If it is false that not every student read any book, it only follows that every student read a book, but it does not follow that there are books that every student read. The same is true for *at most n*, *few* and any other merely DE quantifier. As indicated in (10), sentences in which an EP associates with NPI *any* in the scope of *few* are generally acceptable. This leaves *few* as a problematic case.

The reason for this difference is evidently that the negation of an AA quantifier commutes with existential quantifiers; the negation of a merely DE quantifier does not.

(122) <u>~Q commutes with disjunction iff Q is AA</u>

$Q(A \vee B) \Leftrightarrow Q(A) \wedge Q(B)$	Q is AA
$\sim Q(A \vee B) \Leftrightarrow \sim(Q(A) \wedge Q(B))$	Laws of logic
$\sim Q(A \vee B) \Leftrightarrow \sim Q(A) \vee \sim Q(B)$	DeMorgan's Laws

Given that existential quantification is just generalized disjunction, it follows that the negations of AA quantifiers commute with existential quantification. And, furthermore, if the negation of a quantifier Q commutes with existential quantification, then Q is AA.

So, if we want to distinguish AA quantifiers from merely DE quantifiers in such constructions, we have a basis for the distinction. Falsity of a statement of form (123) – a precondition for using the exceptive – guarantees that there is a least exception set, when Q is AA.

(123) [Q 1[SOME(A-C_F)2[t_t R t_2]]]

If Q is merely DE, there may be a least exception set, but there may not be. So, we might make the guaranteed existence of a least exception set a precondition for the use of a *but*-phrase. This would force us, however, to abandon the efficient explanation of ungrammaticality in terms of systematic contradiction. Statements of the form (123), where Q is merely DE, are not systematic contradictions they are simply not as reliable in providing least exceptions sets.

3.3 PUZZLE 2

In this section we address the issue of the NPI-licensing properties of Q's modified by EPs. In section 3.3.1, we establish the basic licensing properties of Q's modified by EPs in comparison, in particular, to non-monotonic Q's. In section 3.3.2, we give a detailed analysis of the differences in licensing between the closely related *nobody but John* and *only John*. In these two sections, we show that an approach in which Leastness applies non-locally gives the best account of the licensing properties of Q's modified by EPs. Section 3.3.3 concludes.

3.3.1 NPI-licensing properties of [*no* A *but* C]

Any and Exceptives

Now we turn to the solution for our puzzle about NPI-licensing. As noted above, NPIs are licensed in the scope of the surface constituent [no NP but DP]:

(124) a. Nobody but Bill said anything (*any*/*ever*)

b. Nobody but Bill lifted a finger to help (minimizers)

c. Nobody but Bill likes PANCAKES either ('strong' NPIs)

Constituent analyses (von Fintel 1993, Moltmann 1995, Postal 2000) predict that [no NP but DP] has the denotation in (125).

(125) $[\![\text{no A but C}]\!] = \lambda X.X \cap A = C$

This function is non-monotonic. This non-monotonicity is supported by sentence-level judgments of the following kind:

(126) 'nobody but Bill left' does not entail 'nobody but Bill left early'

It is true that non-monotonic functions sometimes license NPIs in their scope, as in (127), though such licensing is generally heavily context dependent, see (128).

(127) Exactly three students said anything in my class.

(128) ??Exactly twelve of my thirteen students said anything in my class.

However, non-monotonic functions are not strong NPI licensers. And in fact, the non-monotonic quantifier [every A but C] does not license any NPIs at all.

(129) a. *Exactly three students left until five.

b. *Exactly three students have visited me in weeks.

c. Exactly three students like pancakes. *Exactly three pancakes like waffles either.

(130) *Every student but Mary ever read anything.

As indicated above, [no A but C] – like [no NP] – is a strong NPI licenser.

(131) a. Nobody but Bill left until five.

b. Nobody but Bill has visited me in weeks.

c. Nobody but Bill likes PANCAKES, either.

(132) a. No student left until five.

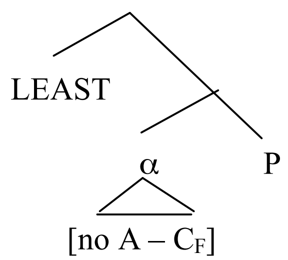
b. No student has visited me in weeks.

c. No student likes PANCAKES, either.

Strong NPI licensing is generally attributed to Anti-Additivity (AA; Zwarts 1998, van der Wouden 1997, a.o.) – a property incompatible with non-monotonicity. For example, [no NP] is AA, [not every NP] is DE but not AA.

The divided approach to the semantics of EPs offers a new perspective on these facts. We have analyzed the semantics of EPs into set subtraction and an operator LEAST which takes clausal scope. Consider again the simple case of a sentence containing [no A but C]. At LF, the sentence will have the structure in (133).

(133)



At LF, LEAST does not take the DP headed by *no* as an argument projecting a non-monotonic generalized quantifier.[16] Instead it takes scope at a node of type t. That leaves a garden variety negative universal in subject position, with an alternative-generating constituent of type <e,t> in its restrictor. This negative universal (constituent α in (133)) is Downward Entailing and, in fact, Anti-Additive. Thus, to explain the occurrence of strong NPIs in the scope of [no A but C] all we need assume is (i) a standard rule that strong NPIs must be c-commanded by an AA function and (ii) our current approach to EPs, which independently helped us assign correct truth conditions to sentences in which an EP associates with NPI *any*.

3.3.2 Only vs. Nobody but

In this section, I argue that constituent analyses of EPs cannot rely on von Fintel's (1999) approach to NPI-licensing to save their analyses. Our discussion of [no A but C] is reminiscent of the case of *only*. *Only* licenses NPIs in its scope, see (134). And sentence-level monotonicity inferences such as (126) also fail for *only*, see (135).

(134) Only Bill *ever* talked to *anyone*.

(135) 'only Bill left' does not entail 'only Bill left early'

Von Fintel 1999 explains the apparently contradictory evidence in (134) and (135) by arguing that, though *only* is not DE in the strict sense, it is Strawson DE, see (137).

(136) $[\![\text{only}]\!](x)(P)$ is defined only if $P(x) = \text{True}$.

If defined, $[\![\text{only}]\!](x)(P) = \text{True}$ iff $\sim\exists y \neq x$: $P(y) = \text{True}$.

(137) Strawson Downward Entailingness

A function f of type $<\sigma,\tau>$ is Strawson-DE iff for all x, y of type σ such that $x \Rightarrow y$ and $f(x)$ is defined: $f(y) \Rightarrow f(x)$.

So, an alternative to the divided-EP explanation of the facts in (124) might be to say that [no A but C], like *only* DP, is Strawson-DE. This can be done simply by changing von Fintel's (1993) denotation of *but* from (138) to (139).

(138) $[\![\text{but}]\!] = \lambda C_{et}.\lambda D_{<et,ett>}.\lambda A_{et}.\lambda P_{et}.P \in D(A-C) \& \forall S(P \in D(A-S) \rightarrow C \subseteq S)$

(139) $[\![\text{but}]\!] = \lambda C_{et}.\lambda D_{<et,ett>}.\lambda A_{et}.\lambda P_{et}{:}\forall S(P \in D(A-S) \rightarrow C \subseteq S).\ P \in D(A-C)$

This gives us the denotation in (140) for the constituent [no A but C] (see proofs in (27)-(28)).

(140) $[\![\text{no A but C}]\!] = \lambda P{:}C \subseteq P \cap A.\, P \cap A \subseteq C$

This makes DPs like *only Bill* and *nobody but Bill* essentially equivalent. It is striking to note then that they differ in their NPI-licensing properties. Consider for example the contrast in (141).

(141) a.Nobody but Bill likes pancakes.

Nobody but Bill likes WAFFLES, either.

b. Only John likes pancakes.

*Only John likes WAFFLES, either. (Nathan 1999)

The following items show a similar contrast:

(142) Nobody but John has arrived **yet**

*Only John has arrived **yet**

(143) ?Nobody but John arrived **until 5**

*Only John arrived **until 5** (cf. Atlas 1993)

(144) Nobody but Sue has visited John **in years**

*Only Sue has visited John **in years**

(145) ?Nobody but Chris **slept a wink** last night

?#Only Chris **slept a wink** last night

Nobody slept a wink last night but Chris (Horn 1996 (17c))

The question is why, if these two quantifiers are semantically equivalent, do they show different licensing properties. This puzzle is sharpened by noticing that *only* DP is not only Strawson DE, but Strawson AA (i.e. is Strawson DE and license inference like (146), see definition of AA in (8)).

(146) Only Bill swims and Only Bill smokes $\Rightarrow$ Only Bill swims or smokes

I propose that the divided-EP analysis is better suited to explain the contrasts in (141) - (145) than a Strawson constituent analysis. Notice that these strong NPIs are also unacceptable in other environments that von Fintel (1999) diagnoses as Strawson DE:

(147) Adversatives

Fred hasn't arrived. *Bill is sorry Sue has arrived either.

*Sue is sorry that Bill has arrived yet

*Sue is sorry that Bill arrived until five

*Sue is sorry that Bill has visited John in years

?Sue is sorry that she slept a wink last night

(148) Antecedent of a Conditional

Fred isn't here. *If Bill is here either, then Mary is upset

??If Bill has arrived yet, then Mary is upset

*If Bill arrived until five, Mary was upset.

*If Sue has visited Bill in years, then Mary is upset.

?If Bill slept a wink last night, Mary was upset.

I propose, then, that there is a class of NPIs that are not satisfied with Strawson-DE (or Strawson-AA, for that matter) licensers, but must instead appear in the LF scope of an operator that is AA defined in terms of standard entailment. In this case, the divided EP analysis gives a neat account of the facts.[17]

3.3.3 Conclusion

In this section, we have seen that a divided-EP theory gives the best account of the NPI-licensing properties of Q's modified by EPs. The most important observation is that [no A but C] licenses strong NPIs, unlike non-monotonic quantifiers and unlike Strawson DE quantifiers. The difference between [no A but C] and [only C] is observed in Atlas 1993, Horn (1996) and Nathan (1999), a.o. The key to our success in these cases has been to allow LEAST to take clausal scope. This use of wide scope possibilities for LEAST raises several questions. Can LEAST freely take scope above any c-commanding operator? Do we find ambiguities in exceptive sentences that depend solely on the scope of LEAST? I discuss these matters in the next section, along with another possible problem.

3.4 PUZZLE 1: PART 2

In part 1 (section 3.2), we assigned the correct truth conditions to sentences involving an EP associated with NPI *any*. Now we need to address the issue of the distribution of EPs. This essentially boils down to explaining the possible scopes for our LEAST operator. Two major issues arise: (Issue 1) do we find scope ambiguities in sentences involving EPs? In other words, in a sentence where two possible scope sites yield non-trivial truth conditions are both possible scope sites for LEAST? (Issue 2) how can we explain the fact that EPs only felicitously associate with *any* in a DE environment and not with any other indefinites that are allowed to occur in DE environments?

3.4.1 ISSUE 1

Given that we posit non-local scope for LEAST when it binds an EP associated with NPI *any*, we need to address the issue of when non-local scope is possible. For example, do we find scope ambiguities involving EPs? The answer appears to be no. Consider (149) and (150).

(149) Every student read no book but War & Peace.

(150) Some student read no book but War & Peace.

In each of these sentences there are two possible scope sites for LEAST that yield consistent truth conditions:

(151) LOCAL SCOPE

a. Every student$_1$ [LEAST [no book – WP$_2$ [t_1 read t_2]]

b. Some student$_1$ [LEAST [no book – WP$_2$ [t_1 read t_2]]

(152) NON-LOCAL SCOPE

a. LEAST [Every student$_1$ [no book – WP$_2$ [t_1 read t_2]]

b. LEAST [some student$_1$ [no book – WP$_2$ [t_1 read t_2]]

(151a) implies that every student read WP and no other book; (151b) that some student read WP and no other book. These correspond to the intuitive readings of (149) and (150), respectively. What of (152a) and (152b)? Do they yield distinct readings? Are these readings available? To answer these questions we can rely on our general result about the truth conditions of sentences involving NPI *any* and non-local scope, i.e., (94). We can lean on this result because the complements of LEAST in (152) are equivalent to structures of the form covered by (94).[18]

(153) a. No student read any book but WP (=(152a))

b. Not every student read any book but WP (=(152b))

(154) (153a) = 1 iff {WP} = {x: x is a book and some student read x}

(153b) = 1 iff not every student read a book that is not WP and

{WP} = {x: x is a book and every student read x}

Are these readings available for (149) and (150)? No. (149) is not compatible with some student not reading W&P. (150) does not require that every student read W&P. This suggests to me that there must be a minimality constraint on the scope of LEAST.

(155) LEAST takes scope at the minimal node of type t dominating the EP it binds.

But wait, this rules out the cases with NPI *any*. Perhaps a better constraint would be (156).

(156) LEAST takes scope at the minimal node of type t dominating the EP it binds, at which it yields non-trivial truth conditions.

This looks more promising, but leads us directly into Issue 2.

3.4.2 ISSUE 2

If (156) is the only constraint on the scope of LEAST, then we predict that (159) should be equivalent to and just as grammatical as (158), given that (157a) and (157b) are equivalent.

(157) a. No student read any book on the list

b. No student read a book on the list

(158) No student read any book on the list but Moby Dick

(159) *No student read a book on the list but Moby Dick

Similarly in environments where *some*'s positive polarity is neutralized we expect it to associate with EPs:

(160) a. You can't tell me Bill didn't read some book

b. *You can't tell me Bill didn't read some book but Moby Dick

I propose that we make use of *any*'s properties as an NPI and LEAST's status as a part of the recursive pragmatics to state an appropriate constraint on the scope of LEAST. Consider the version of Kadmon & Landman's approach to licensing *any* as adapted by Chierchia (2004). In Chierchia's theory, there is an operator **O** that binds the domain variable of *any* and tests it to see if widenings of the domain make the statement stronger. Chierchia stipulates that **O** may only take scope just above DE functions. Suppose we depart from Chierchia in this assumption and think of **O** as a test that is imposed at each node of type t dominating *any* such that *any* is licensed only if it passes the test at some node.

(161) $[\![\text{any}]\!] = \lambda D.\ \lambda A.\ \lambda B.\ (A \cap D) \cap B \neq \varnothing$

(162) *any* marks it domain argument with an F: [*any* D_F]

(163)
$$\mathbf{O}(\langle F,D \rangle) \begin{cases} = F(D), \text{ if } \forall D' \, [D \subseteq D' \rightarrow F(D') \Rightarrow F(D)]^{19} \\ = \langle F,D \rangle, \text{ otherwise} \end{cases}$$

(164) Bill didn't read any_D book

(165)
1 iff $(\text{BOOK} \cap D) \cap \{y: \text{Bill read } y\} = \varnothing$ TEST PASSED
O $\langle \lambda X.\ (\text{BOOK} \cap X) \cap \{y: \text{Bill read } y\} = \varnothing, D \rangle$
Not $\langle \lambda X.\ (\text{BOOK} \cap X) \cap \{y: \text{Bill read } y\} \neq \varnothing, D \rangle$ TEST FAILED
O $\langle \lambda X.\ (\text{BOOK} \cap X) \cap \{y: \text{Bill read } y\} \neq \varnothing, D \rangle$
[Any D_F] book_1
Bill
read t_1

Perhaps, the requirement that *any* pass the test at some node could be modeled as a requirement that the focus denotation of the root node of a sentence always be a 1-tuple.

What I would like to propose is that LEAST always takes the minimal scope possible, where a scope position is impossible if either (i) it is incompatible with LEAST typewise or (ii) it has failed the test **O**. A scope site should be regarded as possible even if its output is trivial truth conditions. This allows us to rule out EPs with existentials that do not have to pass the test **O**, such as *some* or *a*. For these indefinites, the minimal possible scope site will always be the node of type t immediately dominating the existential. That is, non-local scope of LEAST is not a possibility.

The guiding intuition behind this constraint is that the pragmatic operator LEAST is not going to waste its sweetness on the desert air of a node that has been branded a failure by the recursive pragmatics for an independent reason. To try another metaphor, the failure of test **O** is a warning flag the pragmatics sends up to tell LEAST to take a higher scope.

As the reader will have gauged from my waxing metaphorical I do not know precisely how to explain such a constraint. My hope is that it could be derived from independent properties of systems such as Chierchia's 2004 and Fox's 2006 but at the moment I have nothing to offer in this way.

One apparent prediction of this approach is that negative polarity sensitivity should be enough to allow an existential to be compatible with an EP. For example, we might expect minimizers to allow EPs. Take *a single* NP.

(166) ??Bill hasn't read a single book but Moby Dick

??No one read a single book but Moby Dick

The judgments I've gathered from other English speakers is that these sentences are not as good as comparable ones with *any*, but they are improvements over ones with *a* or *one*. Other minimizers such as quantificational superlatives seem to categorically exclude EPs.

(167) *Bill hasn't read the easiest novel but Dick and Jane.

It is unclear what these judgments tell us about our hypothesis for EPs and *any*. After all, it may be that the licensing of *any* and the licensing of minimizers proceed in very different ways in the grammar. Linebarger (1980) and Heim's (1984) discussions of the "iceberg lettuce" examples suffice to show that the licensing principles of *any*/*ever* are interestingly distinct from those of minimizers.

Finally, note the following prediction made by the idea that failing **O** disqualifies a constituent as a scope site for LEAST: an indefinite should be compatible with an EP if the node immediately dominating the indefinite contains an unlicensed *any*. Schematically, we expect EPs to occur in the following environments.

(168) a. DE[…[a NP but DP] […any NP…]]

b. DE[…[a NP[…any NP…] but DP]…]

While such sentences are not perfect, they do seem to be significant improvements over sentences that do not contain NPIs. Consider the following exemplification of (168b).

(169) a. ??No one read a book that criticizes any President but *Against All Enemies*

b. *No one read a book (that criticizes a President) but *Against All Enemies*

A native-speaking anonymous reviewer endorses this prediction and provides the contrast in (170), an exemplification of (168a), as further support.

(170) a. ??No one gave a woman but Susan any money.

(*No one gave a woman but Susan a book.)

b. *No one saw a woman but Susan.

On balance, there is promise in this approach to the distribution of EPs.

4. Conclusion

In this paper, we have presented two puzzles involving the semantics of connected exceptive phrases and shown that existing theories are not able to solve them. We have proposed that a solution to both puzzles can be arrived at through a modification of the compositional implementation of von Fintel's 1993 semantics for exceptive constructions. Specifically, we were able to derive correct truth conditions for sentence in which an EP associates with a DP headed by NPI *any* and to explain the NPI-licensing properties of surface constituents of the form [*no* A *but* C]. The crucial part of our analysis is imposing von Fintel's Leastness condition at scopes non-local to the surface position of the EP. Our analysis faces the problem of licensing EPs with non-NPI existentials in the scope of negation. I suggest that the choice of scope site for Least is sensitive to the output of another strengthening operator, specifically Chierchia's 2004 **O**.

APPENDIX

Proof One

First in this Appendix we prove that the general result given in (94) – namely that for structures like (93), repeated below, the truth conditions in (171) are equivalent to those in (94), when Q is DE. (I abbreviate S∈Q, as Q(S).)

(93) LEAST[Q 1[[SOME [A–C]] 2[t_1 R t_2]]]

(171) $[\![$ (93) $]\!]$ = 1 iff $Q(\{x: \{y: xRy\} \cap A{-}C \neq \varnothing\})$ and

$$\forall S[Q(\{x: \{y: xRy\} \cap A{-}S \neq \varnothing\}) \rightarrow C \subseteq S]$$

(94) $[\![$ (93) $]\!]$ = 1 iff $Q(\{x: \{y: xRy\} \cap A{-}C \neq \varnothing\})$ and

$$C = \{z: z \in A \;\&\; \neg Q(\{y: yRz\})\}$$

Given: I. $Q(\{x: \{y: xRy\} \cap A{-}C \neq \varnothing\})$

II. $\forall S[Q(\{x: \{y: xRy\} \cap A{-}S \neq \varnothing\}) \rightarrow C \subseteq S]$

III. Q is Downward Entailing

IV. $Q(\varnothing)$

Prove that $C = \{z: z \in A \;\&\; \neg Q(\{y: yRz\})\}$

Step 1: Prove that $\{z: z \in A \ \& \ \neg Q(\{y: yRz\})\} \subseteq C$

1. Suppose $\mathbf{a} \in A \ \& \ \neg Q(\{y: yR\mathbf{a}\})$

2. Assumption for reductio: $\mathbf{a} \notin C$

3. Then $\mathbf{a} \in A - C$

4. For arbitrary $\mathbf{b}$, if $\mathbf{b}R\mathbf{a}$, then $\{y: \mathbf{b}Ry\} \cap A - C \neq \varnothing$

5. Therefore $\{x: xR\mathbf{a}\} \subseteq \{x: \{y: xRy\} \cap A - C \neq \varnothing\}$

6. Given III, line 1 and line 5, $\neg Q(\{x: \{y: xRy\} \cap A - C \neq \varnothing\})$

7. But this contradicts I,

8. Therefore $\mathbf{a} \in C$

9. Since $\mathbf{a}$ was arbitrarily chosen we conclude that

$\{z: z \in A \ \& \ \neg Q(\{y: yRz\})\} \subseteq C$

Step 2: Prove that $C \subseteq \{z: z \in A \ \& \ \neg Q(\{y: yRz\})\}$

1. Suppose that $\mathbf{a} \in C$

2. Then, $C \nsubseteq \{\mathbf{a}\}'$

3. Then, given II, $\neg Q(\{x: \{y: xRy\} \cap A - \{\mathbf{a}\}' \neq \varnothing\})$ by contraposition

4. So, $\neg Q(\{x: \{y: xRy\} \cap A \cap \{\mathbf{a}\} \neq \varnothing\})$ (set-theoretic equivalence)

5. I.e., $\neg Q(\{x: xR\mathbf{a} \ \& \ \mathbf{a} \in A\})$

6. Given IV, $\{x: xR\mathbf{a} \ \& \ \mathbf{a} \in A\} \neq \varnothing$

7. Thus, $\mathbf{a} \in A$ and

8. therefore $\{x: xR\mathbf{a} \ \& \ \mathbf{a} \in A\} = \{x: xR\mathbf{a}\}$

9. So, $\neg Q(\{x: xR\mathbf{a}\})$ and $\mathbf{a} \in A$

10. Since **a** was arbitrarily chosen we conclude that

$C \subseteq \{z: z \in A \ \& \ \neg Q(\{y: yRz\})\}$

Thus it is proven from I-IV that

V. $C = \{z: z \in A \ \& \ \neg Q(\{y: yRz\})\}$

And of course it follows from I that I holds. Now we prove I and II follow from I and V, under the assumption of III and IV.

I follows from I

Prove II: (I will prove the contrapositive)

1. Suppose $C \not\subseteq S$
2. Then there is a **a** such that $\mathbf{a} \in C$ and $\mathbf{a} \notin S$
3. That is, by V, $\mathbf{a} \in A$ and $\neg Q(\{y: yR\mathbf{a}\})$ and $\mathbf{a} \notin S$
4. Thus, $\mathbf{a} \in A{-}S$
5. So, $\{y: yR\mathbf{a}\} \subseteq \{x: \{y: xRy\} \cap A{-}S \neq \varnothing\}$ (see lines 3-5 of step 1)
6. Therefore $\neg Q(\{x: \{y: xRy\} \cap A{-}S \neq \varnothing\})$, by III and line 3
7. Since S was arbitrarily chosen we conclude that

$\forall S[Q(\{x: \{y: xRy\} \cap A{-}S \neq \varnothing\}) \rightarrow C \subseteq S]$ (by contraposition)

So we have proven that, under assumptions III and IV,

I. $Q(\{x: \{y: xRy\} \cap A - C \neq \varnothing\})$

II. $\forall S[Q(\{x: \{y: xRy\} \cap A - S \neq \varnothing\}) \rightarrow C \subseteq S]$

Are equivalent to

I. $Q(\{x: \{y: xRy\} \cap A - C \neq \varnothing\})$

V. $C = \{z: z \in A \ \& \ \neg Q(\{y: yRz\})\}$

Proof Two

Suppose on the other hand that Q is UE. Now we prove that I and II are not consistent.

Assume: VI. Q is UE

VII. $C \neq \varnothing$

I. $Q(\{x: \{y: xRy\} \cap A - C \neq \varnothing\})$

II. $\forall S[Q(\{x: \{y: xRy\} \cap A - S \neq \varnothing\}) \rightarrow C \subseteq S]$

Prove that I, II, VI and VII are not consistent.

1. Given VII, there is a set **S** such that $\mathbf{S} \subset C$
2. Therefore, $A - C \subseteq A - \mathbf{S}$
3. Thus, for arbitrary **a**,

   $\{y: \mathbf{a}Ry\} \cap A - C \subseteq \{y: \mathbf{a}Ry\} \cap A - \mathbf{S}$

4. This implies that if $\{y: \mathbf{a}Ry\} \cap A - C \neq \varnothing$, then $\{y: \mathbf{a}Ry\} \cap A - \mathbf{S} \neq \varnothing$

5. Since **a** was arbitrarily chosen we conclude that

$\{x: \{y: xRy\} \cap A - C \neq \varnothing\} \subseteq \{x: \{y: xRy\} \cap A - \mathbf{S} \neq \varnothing\}$

6. Therefore, given I and VI, $Q(\{x: \{y: xRy\} \cap A - \mathbf{S} \neq \varnothing\})$

7. However, since $\mathbf{S} \subset C$ (line 1), $C \nsubseteq \mathbf{S}$

8. Therefore, by II and contraposition, $\neg\ Q(\{x: \{y: xRy\} \cap A - \mathbf{S} \neq \varnothing\})$

9. Line 6 and Line 8 are contradictory so I, II, VI and VII are not consistent

---

[1] Connected exceptive phrases, as in (1), have roughly the syntax of relative clauses, appearing either post-nominally or right-extraposed. These show different properties from free exceptive phrases, cf. (i), which have roughly the syntax of sentence adverbs:

(i) Besides/apart from/except for Bill, every student came.

For discussion of the differences between these two constructions see Hoeksema 1987, Hoeksema 1995 von Fintel 1993.

[2] Von Fintel (1993) refers to this as Uniqueness or unique minimality.

[3] Strictly speaking, this is not true. ↑mon quantifiers actually *guarantee* the existence of a minimal exception: the empty set. To be precise then, we should say that ↑mon quantifiers never have non-empty least exceptions.

[4] Alternatively, these could be analyzed as cases of ellipsis, see Hoeksema 1995 for discussion:

(i) All students but a most five students came.

---

[5] I will not dedicate a section of this review to Szabolcsi 2004. That paper, as I understand it, must make use of Moltmann's semantics for exceptive since it analyzes EPs with NPI *any* in terms of dyadic quantification.

[6] In this paper, I give only an informal sketch of Moltmann's theory. I refer the reader to Moltmann (1995) for details.

[7] Moltmann suggests that (i) receive the denotation in (ii)

(i) but Mary [no man, any woman]

(ii) $\bigcup_{R' \in \{X' \times \{Mary\} : \exists X(X' \subseteq X \ \&\ X \in W(\text{NO MAN}))\}} \{R \backslash R' : R \in (\text{NO MAN, ANY WOMAN})\}$

One easy to fix problem with this is that Moltmann must mean "R∪R'" rather than "R\R'" in her description of the set. There is a more serious problem with the specification of the relations being quantified over by ∪. Moltmann says that these are the Cartesian products of subsets of the witness set of NO MAN with {Mary}. The problem is that the witness set of NO MAN is ∅. (Moltmann makes the confusing claim that a subset of the witness set of NO MAN is "a set containing at least one man" (p.273).) ∅×{Mary} is just ∅; so the set of relations being quantified over is {∅}. This has the effect that BUT MARY(NO MAN, ANY WOMAN) = <NO MAN, ANY WOMAN>. A possible patch would be to assume that Moltmann means not the witness set of NO MAN but the smallest live-on set, namely MAN. This is nearly correct, except that ∅ is a subset of MAN. Including ∅ makes the prediction that (i) is compatible with no man dancing with Mary.

(i) No man danced with any woman but Mary

This is not correct. (i) clearly entails that some man danced with Mary. Perhaps, a restriction could be added to non-empty subsets. At this point, however, I have no idea if this represents Moltmann's intentions.

[8] Note that *but* is not as acceptable with binary exceptions as *except*:

(i) ?Every woman danced with every man but Sue with Bill.

[9] Von Fintel (class notes, 2000) moves EPs to the edge of DP to account for cases like (i).

(i) Every president's wife but Hillary Clinton (Moltmann 1995)

[10] One crucial difference between our proposed operator and these other pragmatic operators is that its contribution is not defeasible. In this way it resembles more the operators Chierchia (2006) uses with dedicated free choice items in Italian, like *qualsiasi*.

[11] In previous work (Gajewski 2004), I suggested that the EP itself takes scope at LF. I am now convinced this cannot be. It is well known that NPI *any* may be licensed across syntactic islands. EPs may associate with *any* even in these environments. If the EP had to scope above the licenser at LF, it would have to cross the island:

(i) a. Sue didn't read [anyone but Bill]'s paper — Left Branch
b. Sue doesn't want cake or [any liquor but gin]. — Coordinate Structure

[12] Actually the picture is more complicated. The semantics proposed does not restrict attention to subsets of the restrictor; so in fact the greatest exception set is U – (A∩P), where U is the universe of discourse. Restriction to subsets of the restrictor is natural and could be accomplished with a presupposition or covert domain restriction.

[13] Thanks to an anonymous reviewer for this example.

---

[14] Postal 2005 reports the judgments in (i).

(i) a. No woman said anything (except hello).

b. Not many women said anything (*except hello).

c. Not every woman said anything (*except hello).

d. Not more than three women said anything (*except hello).

e. Exactly three women said anything (*except hello).

f. If those women said anything (*except hello), it is unusual.

g. Should those women say anything (*except hello)?

Though I would assign a question mark (or two) to (ic), (id) and (ie), I find the others acceptable.

[15] See García-Alvarez (2003) for a claim that EPs have a wider distribution than commonly assumed.

[16] The astute reader will have noticed that a pointwise extension of our operator LEAST to all conjoinable types would allow us to generate just such a structure.

[17] A potentially problematic case is superlatives, which are Strawson DE and not AA as defined by standard entailment. Superlatives at least license some strong NPIs:

(ii) Erin is the tallest girl Fred has ever seen.

?She is the tallest girl Bill has ever seen either.

Erin is the tallest girl John has seen yet.

?Erin was the tallest girl John saw until 5.

?Erin was the tallest girl John (even) lifted a finger to help.

To maintain our analysis we must find a structure for superlatives that contains a strictly AA operator and another operator that masks its monotonicity. I leave this for future research.

[18] A reviewer observes that it is difficult to get a wide scope strengthening for *other than* in these environments as well. For example, it is hard to read (ia) as equivalent to (ib).

(i) a. Some student read no book other than W&P ?⇔?

b. Not every student read any book other than W&P

This is true; though as we have observed in section 3.2.2 there are scope differences between *but* and *other than*. So, (i) suggest that *other than* is also constrained in the scope sites for strengthening.

[19] To this point I have kept the semantics completely extensional, to impose this condition we would just need a straightforward intensionalization of the semantics.

REFERENCES

Atlas, J.: 1993, 'The Importance of Being *Only*: Testing the Neo-Gricean Versus Neo-Entailment Paradigms', *Journal of Semantics* **10**, 301-318.

Baker, C.L.: 1970, 'Double Negatives', *Linguistic Inquiry* **1**, 169-186.

Barwise, J. and R. Cooper: 1981, 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* **4**, 159-219.

Carlson, G.: 1981, 'Polarity *any* is Existential', *Linguistic Inquiry* **11**, 799-804.

Chierchia, G.: 2004, 'Scalar Implicatures, Polarity Phenomena and the Syntax/Pragmatics Interface', in A. Belletti (ed.), *Structures and Beyond*. Oxford University Press.

Chierchia, G.: 2006, 'Broaden Your Views: Implicatures of Domain Widening', *Linguistic Inquiry* **37:4**, 535-590.

von Fintel, K.: 1993, 'Exceptive Constructions', *Natural Language Semantics* **1**, 123-148.

von Fintel, K.: 1994, *Restrictions on Quantifier Domains*, PhD dissertation, University of Massachusetts, Amherst.

von Fintel, K.: 1999, 'NPI Licensing, Strawson-Entailment, and Context-Dependency', *Journal of Semantics* **16**, 97—148

von Fintel, K.: 2000, 'Exceptive Constructions', Class Notes MIT

Fox, D.: 2006, 'Free Choice and the theory of Scalar Implicatures', ms., MIT

Gajewski, J.: 2004, 'Raising Exceptions', in V. Chand, A. Kelleher, A. Rodríguez, and B. Schmeiser (eds.), Proceedings of the 23rd West Coast Conference on Formal Linguistics, pp. 209-222.

Garcia Alvarez, I.: 2003, 'Quantifiers in Exceptive NPs', in G. Garding and M. Tsujimura (eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, pp. 207-216.

Giannakidou, A.: (2006), '*Only*, emotive factives and the dual nature of polarity dependency' *Language* **82:3**, 575-603.

Heim, I.: 1984, 'A Note on Negative Polarity and Downward Entailingness', in C Jones and P. Sells (eds). *Proceedings of NELS 14*, pp. 98-107.

Heim, I. 1987, 'Where does the Definiteness Effect Apply?' in E. Reuland and A. ter Meulen, *The Representation of (In)definiteness*, pp.21-42. MIT Press, Cambridge.

Hoeksema, J.: 1987, 'The Logic of Exception', *ESCOL 1987*, pp. 100-113.

Hoeksema, J.: 1995, 'The Semantics of Exception Phrases' in J. van der Does and J. van Eijk (eds.), *Quantifiers, Logic and Language,* pp. 145-177. CSLI, Stanford.

Horn, L.: 1972, *On the Semantic Properties of the Logical Operators in English*, PhD dissertation. University of California, Los Angeles.

Horn, L.: 1989, *A Natural History of Negation*, University of Chicago Press.

Horn, L.: 1996. 'Exclusive Company: *Only* and the Dynamics of Vertical Inference', *Journal of Semantics* **13**,1-40.

Jacobs, J.: 1983, Fokus und Skalen, zur Syntax and Semantik von Gradenpartikeln im Deutschen. Tübingen: Niemeyer.

Kadmon, N.: 1987, *On Unique and Non-unique Reference and Asymmetric Quantification*, PhD dissertation, University of Massachusetts, Amherst.

Kadmon, N. and F. Landman: 1993, 'ANY', *Linguistics and Philosophy* **16**, 353-422.

Keenan, E. and J. Stavi: 1986, 'A Semantic Characterization of Natural Language Determiners', *Linguistics and Philosophy* **9**, 253-326.

Krifka, M.: 1991, 'A Compositional Semantics for Multiple Focus Constructions' in S. Moore and A.Z. Wyner, *Proceedings of SALT I*, pp. 127-158.

Ladusaw, W.: 1979, *Polarity Sensitivity as Inherent Scope Relations*, Ph.D. dissertation, University of Texas, Austin.

Lasnik, H.: 1972, *Analyses of Negation in English*, PhD dissertation, MIT.

Linebarger, M.: 1980, *The Grammar of Negative Polarity*, PhD dissertation, MIT.

Moltmann, F.: 1995, 'Exception Sentences and Polyadic Quantification', *Linguistics and Philosophy* **18**, 223-280.

Nathan, L.: 1999, 'Either: Negative Polarity meets Focus Sensitivity', ms., MIT

Postal, P.: 2000, 'The Ohio Lectures on *Squat*', ms., New York University.

Postal, P.: 2005, 'Suppose (If Only for an Hour) That Negative Polarity Items are Negation-Containing Phrases' handout from Polarity From Different Perspectives Workshop, March 11-13, New York University.

Quine, W.: 1960, *Word and Object*, MIT Press, Cambridge.

Rullmann, H.: 2003, 'Additive particles and Polarity', *Journal of Semantics* **20**, 329-401.

Szabolcsi, A.: 2004, 'Positive Polarity – Negative Polarity', *Natural Language and Linguistic Theory* **22**, 409-452.

van der Wouden, T.: 1997, Negative Contexts: Collocation, Polarity and Multiple Negation. Routledge

Zwarts, F.: 1998, 'Three Types of Polarity', in F. Hamm and E. Hinrichs (eds.), *Plurality and Quantification*, pp. 177-238.