A note on connected exceptive phrases*

Luka Crnič

October 4, 2016

Abstract

Connected exceptive phrases may occur only in certain types of quantified sentences. This has been captured by assigning them meanings that indirectly regulate their distribution: only in combination with certain quantifiers do they yield consistent interpretations (e.g., von Fintel 1993, Moltmann 1995). This paper provides new arguments for splitting the semantic contribution of (connected) exceptives into two structurally separable components, that is, into two components that may be computed at different scope sites in the sentence (esp. Gajewski 2008, 2013). Furthermore, we argue that these components do not stand in a binding/movement dependency relation. We conclude the paper by extending the reasoning to *almost*.

1 Exceptives and their semantics

The distribution of connected exceptive phrases like *but 'War and Peace'* and *except Mary* (henceforth, CEPs) is curiously constrained. For instance, while they may modify (negative) universal quantifiers, they may not modify plain existential quantifiers:

(1) a. Every book but 'War and Peace' is worth reading.
b. No book but 'War and Peace' is worth reading.
c. #Some book but 'War and Peace' is worth reading.

Semantics of exception sentences. It has been proposed that the difference in the acceptability of the sentences in (1) springs from their different truth-conditions, which crucially depend on the semantic contribution of the CEPs that occur in them (e.g., von Fintel 1993, Moltmann 1995). More to the point, whereas the truth-conditions of (1a)-(1b) are contingent, the truth-conditions of (1c) are contradictory (as well as those of all its variants in which non-functional vocabulary is replaced by some other material). The assumption is then that when a sentence is contradictory in this way (independently of what non-functional material is used in it), it is judged to be ungrammatical (see esp. Gajewski 2002, Chierchia 2013 for discussion). We spell out this approach to CEPs more fully in the remainder of the section, taking von Fintel's (1993) theory as our template.

On von Fintel's (1993) theory, the truth-conditions of sentences with CEPs are those schematized in (2), where X is the singleton set containing the excepted element ('War and Peace' in the above examples) and E stands for the domain of individuals.

*Acknowledgements to be added.

(2) $[\![\text{D P [but X] Q}]\!] = 1$ iff $\underbrace{\text{D(P}\backslash\text{X)(Q)}}_{\text{Domain Subtraction}} \wedge \underbrace{\forall \text{X'}\subseteq\text{E: D(P}\backslash\text{X')(Q)} \rightarrow (\text{X} \subseteq \text{X'})}_{\text{Uniqueness}}$

These truth-conditions consist of two components – Domain Subtraction and Uniqueness – that conspire to yield the restricted distribution of CEPs. The import of Domain Subtraction is to shift the quantification in the sentence to one in which the excepted element is subtracted from the domain of the quantifier containing the CEP. The import of Uniqueness, which is key in governing the distribution of CEPs, is that the excepted element is the minimal element such that it can be subtracted and the respective quantification yield a true meaning (in other words, subtracting some other set of individuals, which includes the empty set, from the respective restrictor of the quantifier yields either a false meaning or a meaning that is entailed by the meaning that we obtain by subtracting the excepted element).

Grammatical examples. The sentences in (1a)-(1b) have the truth-conditions in (3)-(4), respectively: Domain Subtraction leads to the quantification being true only if the excepted element is subtracted from the restrictor of the quantifier; Uniqueness corresponds to the excepted element, 'War and Peace', forming the minimal set such that subtracting it from the restrictor of the quantifier results in a true quantification. (Semantically equivalent formulas are provided in the parantheses at the bottom of the examples. See von Fintel 1993, Gajewski 2008 for detailed computations.)

(3) $\underbrace{\{\text{x} \mid \text{x is a book}\}\backslash\{\text{WP}\} \subseteq \{\text{x: x is worth reading)}\}}_{\text{Domain Subtraction: Every book that is not WP is worth reading}} \wedge$

$$\underbrace{\forall \text{X}\subseteq\text{E: } (\{\text{x} \mid \text{x is a book}\}\backslash\text{X} \subseteq \{\text{x: x is worth reading}\} \rightarrow \{\text{WP}\} \subseteq \text{X}}_{\text{Uniqueness: \{WP\} is the minimal set X s.t. every book that is not in X is worth reading}}$$

$$\Big(= \{\text{x} \mid \text{x is a book}\} \cap \{\text{x} \mid \neg(\text{x is worth reading})\} = \{\text{WP}\}\Big)$$

(4) $\underbrace{\{\text{x} \mid \text{x is a book}\}\backslash\{\text{WP}\} \cap \{\text{x: x is worth reading)}\} = \emptyset}_{\text{Domain Subtraction: No book that is not WP is worth reading}} \wedge$

$$\underbrace{\forall \text{X}\subseteq\text{E: } (\{\text{x} \mid \text{x is a book}\}\backslash\text{X} \cap \{\text{x: x is worth reading}\} = \emptyset) \rightarrow \{\text{WP}\} \subseteq \text{X}}_{\text{Uniqueness: \{WP\} is the minimal set X s.t. no book that is not in X is worth reading}}$$

$$\Big(= \{\text{x} \mid \text{x is a book}\} \cap \{\text{x} \mid \text{x is worth reading}\} = \{\text{WP}\}\Big)$$

Some inferences that follow from (3)-(4) are provided in (5)-(6): the (a) inferences correspond to the import of Domain Subtraction, while the (b) and (c) inferences follow from (a) conjoined with the Uniqueness inference.

(5) *Inferences following from (3):*
   a. Every book that is not 'War and Peace' is worth reading.
   b. 'War and Peace' is a book.
   c. 'War and Peace' is not worth reading.

(6) *Inferences following from (4):*
   a. No book that is not 'War and Peace' is worth reading.
   b. 'War and Peace' is a book.
   c. 'War and Peace' is worth reading.

To convince ourselves that (5)-(6) are indeed entailed by the meanings above, let us briefly zoom in on (5): on the one hand, if 'War and Peace' is a book, and the singleton set containing 'War and Peace' is the minimal exception set to the quantification, then 'War and Peace' cannot be among the books worth reading – if it were, there would be a smaller exception set that could be subtracted from the restrictor of the quantifier and the quantification still yield a true meaning, namely, the empty set. On the other hand, if 'War and Peace' is not a book, then there is again a smaller exception set that would make the quantification true, namely, the empty set.

Ungrammatical examples. Sentence (1c) has a contradictory meaning on this analysis: it cannot hold that some book other than 'War and Peace' is worth reading and that {'War and Peace'} is the minimal set such that you can subtract it from the restrictor of the quantifier and obtain a true meaning.

(7) $\underbrace{\{x \mid x \text{ is a book}\}\backslash\{WP\} \cap \{x\text{: } x \text{ is worth reading}\} \neq \emptyset}_{\text{Domain Subtraction: Some book that is not WP is worth reading}} \wedge$

$\underbrace{\forall X \subseteq E\text{: } (\{x \mid x \text{ is a book}\}\backslash X \cap \{x\text{: } x \text{ is worth reading}\} \neq \emptyset) \rightarrow \{WP\} \subseteq X}_{\text{Uniqueness: \{WP\} is the minimal set X s.t. some book that is not in X is worth reading}}$

Namely, if you subtract the empty set from the restrictor of the quantifier, which is a proper subset of {'War and Peace'}, you obtain a meaning that will be true whenever the subtraction of {'War and Peace'} leads to a true meaning (and in other cases as well).

Compositionality. How are the truth-conditions described above derived compositionally? As may be most natural given the surface form of the sentence, von Fintel (1993) proposes that both meaning components (Domain Subtraction and Uniqueness) are encoded in the CEP, which is interpreted *in situ*. In other words, the mechanism governing the distribution of CEPs is taken to be part of their meaning.

The meaning of CEP *but 'War and Peace'* is provided in (8). It takes three arguments: a nominal predicate, a quantificational determiner, and the main predicate.

(8) $[\![\text{but 'War and Peace'}]\!] =$

$\lambda P.\ \lambda Q.\ \lambda R.\ Q(P\backslash\{WP\})(R)\ \&\ \forall X \subseteq E\text{: } Q(P\backslash X)(R) \rightarrow \{WP\} \subseteq X$

It is clear that the interpretation of the structures in (9) yields precisely the meanings provided in (3), (4) and (7), respectively. A representative computation of the truth-conditions of (9a) is provided in (10), where the final output corresponds to (3).

(9) a. [every [book [but WP]]] [is worth reading]
b. [no [book [but WP]]] [is worth reading]
c. [some [book [but WP]]] [is worth reading]

(10) $\big([\![\text{but 'War and Peace'}]\!]([\![\text{book}]\!])\big)([\![\text{every}]\!])\big)([\![\text{worth reading}]\!]) = 1$ iff

$[\![\text{every}]\!]([\![\text{book}]\!]\backslash\{WP\})([\![\text{worth reading}]\!]) = 1 \wedge$
$\forall X \subseteq E\text{: } [\![\text{every}]\!]([\![\text{book}]\!]\backslash X)([\![\text{worth reading}]\!]) = 1 \rightarrow \{WP\} \subseteq X$ iff

$\{x \mid x \text{ is a book}\}\backslash\{WP\} \subseteq \{x\text{: } x \text{ is worth reading})\} \wedge$
$\forall X \subseteq E\text{: } (\{x \mid x \text{ is a book}\}\backslash X \subseteq \{x\text{: } x \text{ is worth reading}\}) \rightarrow \{WP\} \subseteq X$

Summary. In von Fintel's (1993) proposal, we have a fully compositional analysis of CEPs that accounts for the data exemplified in (1). The main idea behind the proposal is that the felicity of CEPs depends on the interpretation of the sentences in which they occur, to which their semantic contribution is critical. More to the point, von Fintel assigns CEPs complex meanings that may felicitously combine only with certain quantificational determiners. These meanings consists of two components, Domain Subtraction and Uniqueness, as outlined in (11). (A similar state of affairs obtains also in the proposal of Moltmann 1995, which we cannot discuss here for reasons of space.)

(11) $[_S$ ... $[_{QP}$ every/no/some/etc. [NP [but NP']]] ...]
↓
Subtraction+Uniqueness

We argue in the following section that the distribution of CEPs in ellipsis contexts is incompatible with such an analysis. Thus, while von Fintel (1993) may well be correct that the distribution of CEPs should be explained in semantic terms, and about the meanings that he assigns to the exception sentences in (1), the semantic contribution of CEPs cannot be quite the one he proposed. More specifically, we conclude that whereas Domain Subtraction should be encoded in the meaning of CEPs, Uniqueness cannot be. And so the function of the latter must be picked up by a different mechanism. We show in the conclusion of the paper that the data are compatible with, and thus provide support for, Gajewski's (2013) proposal (but not Gajewski's 2008 proposal; see also Hirsch 2016).

2 Ellipsis puzzles

We present new arguments against encoding both Domain Subtraction and Uniqueness into the meaning of CEPs. The arguments build on the distribution of CEPs in ellipsis contexts. Roughly, in ellipsis contexts, a DP modified by a CEP may antecede a DP modified by an exceptive that is not subject to the same distributional restrictions and does not necessarily give rise to the same inferences. We begin the section by spelling out some background assumptions about ellipsis licensing, which we take to be uncontroversial.

2.1 Conditions on ellipsis licensing

Parallelism. Ellipsis provides a good diagnostic tool for determining the structure as well as the semantics of different types of phrases (see, e.g., Fox 2000, Hartman 2011 for recent applications). This holds because ellipsis is licensed only if some constituent containing the elided material stands in an identity relation to some antecedent constituent. Although the precise nature of this identity relation as well as the necessary size of constituents involved is still under debate, a minimal condition on ellipsis licensing – be it IP, VP, or NP ellipsis – is that the meaning of the antecedent constituent must be in the focus value of the constituent containing the elided material, as stated in (12) (cf. Rooth 1992, Heim 1996, Fox 2000, among others; Elbourne 2013 for NP ellipsis).[1]

[1]The focus value of a constituent is computed as in (i) (following Rooth 1985).

(i) a. If X is a terminal node that is not F-marked, then ALT(X) = {⟦X⟧}
b. If X is a terminal node that is F-marked, then ALT(X) = {⟦Y⟧ | type(Y) = type(X)}
c. If X = [Y Z] is a branching node such that the meaning of Z is an argument of the meaning of Y, then ALT(X) = {Y'(Z') | Y' ∈ ALT(Y) & Z' ∈ ALT(Z)}

(12) *Parallelism Condition on Ellipsis Licensing*:
Ellipsis of a constituent ϵ is licensed only if at LF ϵ is reflexively dominated by some constituent, β, such that there is an antecedent constituent in the discourse, α, such that the meaning of α is in the focus value of β. (We call α and β the antecedent and the ellipsis Parallelism Domain, respectively.)

While this condition by itself allows for some differences in the morphology of the antecedent and the elided constituents, no difference in semantics is admitted (see van Craenenbroeck & Merchant 2013 for a recent discussion of admitted alternations, and Crnič 2016 for a qualification of the claims about polarity items). Sag's (1976) example with a negative polarity item (NPI) *any* in the antecedent VP is provided in (13): whereas the NPI is licensed in the antecedent VP, it would not be licensed in the elided VP. Accordingly, a semantically equivalent plain existential quantifier must be taken to occur in its parallel position in the elided VP, as given in (14).

(13) a. John didn't read any book. Bill did △.
b. △ = read a book

(14) *Parallelism Condition is satisfied*:
$[\![$[neg [[any book] [λ1 [J. read t_1]]]]$]\!] \in$ ALT([did$_F$ [[a book] [λ1 [B.$_F$ read t_1]]]])
$\big(= \{\hat{}(\text{X read a book}), \hat{}\neg(\text{X read a book}) \mid \text{X} \in D_e\}\big)$

Scope configurations. It turns out that the condition on ellipsis licensing in (12) is too weak to fully capture ellipsis behavior on its own. For example, consider (15).

(15) John solved exactly three (of the ten) exercises. You had to △ to get an A.

The second sentence in (15) is unambiguous: it only conveys that the requirement was that you solve exactly three exercises. This is an odd requirement in natural contexts in which you are usually required to solve as many exercises as possible to get a good grade. The sentence cannot describe solving three exercises as the minimal requirement for getting an A, which is more plausible in natural contexts (say, on a particularly hard exam, solving three exercises may put you in the top 5% of the students, and thus suffice for an A; solving more exercises is, of course, not harmful even when the exams are hard).[2]

In this, the sentence differs from its fully pronounced counterpart in (16). This pragmatically more plausible reading could be obtained by interpreting *exactly three exercises* as taking scope above the universal modal at LF (or perhaps by interpreting *exactly (three)* as taking such scope, cf. Abels & Martí 2010, Beck 2012). (Fiengo & May 1994 and Fox 2000, among others, discuss many further examples like this.)

(16) You had to solve exactly three exercises to get an A.

---

[2]Note that the LFs required for this more plausible interpretation of (15) would satisfy Parallelism, as indicated in (i)-(ii). We underline the Parallelism Domains selected to satisfy Parallelism.

(i) a. [exactly three exercises] [λx [I solved x]]
b. [exactly three exercises] [λy [$\Box_F$ [you$_F$ solved y]]]

(ii) *Parallelism Condition is satisfied*:
$[\![$[exactly 3 exercises] [λx [I solved x]]$]\!] \in$ ALT([exactly 3 exercises] [λy [$\Box_F$ [you$_F$ solved y]]])

This constraint on scope in ellipsis contexts is captured by the descriptive generalization in (17). The constraint states that if the elided constituent contains a quantificational element at surface form, this quantificational element may move above another scope-bearing element at LF, and thus expand the ellipsis Parallelism Domain (henceforth, PD), only if a parallel scope configuration obtains in the antecedent PD (though some of the quantificational elements may differ between the two PDs).[3]

(17) *Scope Parallelism Generalization* (Fox 2000, Ch. 2):
The scope relations between quantificational elements in the ellipsis PD must be identical to those between their anteceding elements in the antecedent PD.

Accordingly, in (15), for *exactly three exercises* (or perhaps just *exactly (three)*) to take scope above the universal modal in the ellipsis PD, a comparable scope configuration would have to obtain in the antecedent PD. Since this is not possible – namely, there are no scope-bearing elements in the first sentence other than *exactly three exercises*, given in (18a) – the second sentence can only be interpreted with *exactly three exercises* taking scope below the modal, as given in (18b). This results in a pragmatically marked meaning. (We underline the PDs selected for Parallelism.)

(18) a. [exactly three exercises] [λx [John solved x]]

b. [□ [[exactly three exercises] [λy [you$_F$ solved y]]]]
(→ Parallelism is satisfied, PD respects (17))

c. *[exactly three exercises] [λy [□$_F$ [you$_F$ solved y]]]
(→ Parallelism is satisfied, but PD violates (17))

Movement from the ellipsis site. A related constraint is that if an element in the elided VP moves at LF and ends up not being (able to be) contained in a PD that satisfies Parallelism, then it needs to be pronounced and bear appropriate focus (see, esp., Heim 1996 for discussion, and Schwarzschild 1999 for a more general take on this issue).

(19) *Moved NP Constraint*:
If an NP moves from an elided constituent at LF but lacks a semantically equivalent element in a parallel position in the antecedent clause, it must be pronounced.

We illustrate the Moved NP Constraint on the basis of the example in (20).

(20) a. John read some novel. Bill did △ too.
b. △ = read some novel, △ ≠ read some book

The second sentence in (20a) cannot have the meaning that Bill read some book. While PDs that would allow for such a meaning and satisfy Parallelism can be constructed, as indicated in (21), *some book* does not satisfy Moved NP Constraint. Thus, the sequence in (20a) is marked on the resolution of the elided VP to *read some book*.

(21) a. [some novel] λ1 [John read t$_1$]

[3]There have been various proposals how to derive this generalization, invariably resulting in a stronger ellipsis licensing condition than the one we adopt in the main text, see, e.g., Fiengo & May 1994, Heim 1996, Fox 2000, Griffiths & Lipták 2014. Instead of adopting one of these more involved treatments of ellipsis, we will simply fall back on (17) in the following.

b. [some book] $\underline{\lambda 2\ [\text{Bill}_{\text{F}}\ \text{read}\ t_2]}$
c. $[\![\lambda 1\ [\text{John read}\ t_1]]\!] \in \text{ALT}([\lambda 2\ [\text{Bill}_{\text{F}}\ \text{read}\ t_2]])$

Again, we refer the reader to Fiengo & May (1994), Heim (1996), Fox (2000), Griffiths & Lipták (2014), among others, for detailed studies of Scope Parallelism Generalization and Moved NP Constraint, as well as for their potential explanations. The precise mechanisms governing these generalizations are not important to us here – what is important is that they should, all else being equal, hold just as much of DPs modified by connected exceptives as they hold of other DPs. We now proceed to the three ellipsis puzzles.

2.2 Existential quantifiers

Consider the sequences in (22)-(24). (The latter two sequences are modeled after examples in Johnson 2001 and Sag 1976, respectively.)[4,5]

(22) John read no book but 'War and Peace'. Mary did △ however.

(23) I could find no solution except to use covert exhaustification, but Irene might △.

(24) Although John will trust no one but the President, Bill will △.

We focus on the first sequence in the following. There, the meaning of the second sentence corresponds to the paraphrase in (25), where an *other than* exceptive is modifying an existential quantifier.

(25) Mary read some book other than 'War and Peace'.

Puzzle. The felicity of the sequence in (22) is unexpected on the approach to CEPs introduced above, on which Domain Subtraction and Uniqueness are encoded in the meaning of *but*. The structure and the meaning of the sentence containing the antecedent VP in (22) on that approach are those provided in (26) (see the discussion of (4) above).

(26) a. John read no book but 'War and Peace'.
b. [no book but WP] [λx [John read x]]
c. {x: x is a book} ∩ {x: John read x} = {WP}

The sentence containing the elided VP in (22) could be assigned one of the representations in (27), in which *Mary* and *did* are focused: the first representation contains a CEP; the second representation lacks an exceptive; and the third representation contains an exceptive distinct from *but 'War and Peace'*, namely, an *other than* exceptive that can modify existential quantifiers. None of these representations forms a licit sequence with the sentence in (26) however, which is why the felicity of (22) is unexpected.

[4]Some speakers find these examples, and in fact all examples in which a negative quantifier antecedes a plain existential quantifier, marked. See van Craenenbroeck & Temmerman 2015 for discussion.

[5]The argument that we make in this subsection can be made just as well with a sequence in which the negative quantifer in the antecedent VP is replaced by an NPI, as in (i). Some speakers prefer this example to (22). However, since this sequence poses independent problems for von Fintel's proposal, as discussed by Gajewski 2008, we stick to the examples with negative quantifiers.

(i) John didn't read any book but 'War and Peace'. Mary did △ however.

(27) a. Mary did △ however.
b. [did$_{F}$ [a book but WP] [λy [Mary$_{F}$ read y]]]
c. [did$_{F}$ [a book] [λy [Mary$_{F}$ read y]]]
d. [did$_{F}$ [a book other than WP] [λy [Mary$_{F}$ read y]]]

First: it is clear that none of the proper subconstituents of the structures in (27) satisfy Parallelism and Moved NP Constraint (note that the indices in (26)-(27) are distinct due to No Meaningless Coindexation, see Heim 1996). For example, consider the representations in (28a)-(28b), where the λ-prefixed constituents are taken to constitute the PDs. While these PDs satisfy Parallelism, as indicated in (28c), the second structure does not satisfy Moved NP Constraint.

(28) a. [no book but WP] [λx [John read x]]
b. [did$_{F}$ [a book] [λy [Mary$_{F}$ read y]]
c. ⟦[λx [John read x]]⟧ ∈ ALT([λy [Mary$_{F}$ read y]])

Second: the clause in (27b), in which a CEP is modifying an existential quantifier, has a contradictory meaning, as discussed in the preceding section. Third: in contrast to (27b), the meanings of the clauses in (27c)-(27d) are licit. Their focus alternatives are those given in (29)-(30), where *other than NP* picks out Domain Subtraction.

(29) ALT([did$_{F}$ [a book] [λy [Mary$_{F}$ read y]]]) =
{^(X read a book), ^¬(X read a book) | X ∈ D$_{e}$}

(30) ALT([did$_{F}$ [a book other than WP] [λy [Mary$_{F}$ read y]]]) =
{^(X read a book that is not WP), ^¬(X read a book that is not WP) | X ∈ D$_{e}$}

However, neither of these sets of alternatives includes the meaning of the antecedent PD that we computed in (26). This is unsurprising given that Uniqueness is not induced in either structure. Parallelism is thus predicted to be violated, and so sequence (22) should be marked, contrary to fact.

We are forced to conclude that an analysis of CEPs that encodes the mechanism governing their distribution – in this case, Uniqueness – into the meaning of CEPs cannot capture the felicity and import of sequences (22)-(24), at least not while maintaining the minimal and uncontroversial assumptions about ellipsis licensing introduced above.

2.3 Universal quantifiers

Consider sequence (31).

(31) In the exam, I solved every exercise but the last one.
To get an A, you really had to △.

The second sentence in the sequence conveys that you were required to solve every exercise that is not the last one to get an A; with respect to the last exercise, it neither requires you to solve it nor does it require you not to solve it. That is, the sequence is felicitous in contexts in which solving the last exercise is not only not detrimental to getting an A, but may even increase your chances of doing so.

Puzzle. This is unexpected on the approach to CEPs introduced above. On that approach, namely, the first sentence of (31) has the representation in (32).

(32) [every exercise but the last one] [λx [I solved x]]

The second sentence can have one of the representations in (33). In the first structure, the quantifier is interpreted in the embedded clause, and the embedded clause is taken to be the ellipsis PD. In the second structure, the quantifier scopes above the modal, and the matrix clause is taken to be the ellipsis PD.

(33) a. [$\Box_F$ [every exercise but the last one] [λy [you$_F$ solved y]]]
b. [every exercise but the last one] [λy [$\Box_F$ [you$_F$ solved y]]]

Neither representation in (33) would lead to sequence (31) being felicitous in natural contexts. On the one hand, the representation in (33a) has a meaning that the second sentence in (31) clearly need not convey: it entails that you were required not to solve the last exercise in order to get an A, as represented in (34).

(34) $[\![(33a)]\!] = 1$ iff $\Box(\hat{}\{x\text{: you didn't solve x}\} \cap \{x\text{: x is an exercise}\} = \{L\})$

On the other hand, although the representation in (33b) would yield the observed interpretation of (31), it violates Scope Parallelism Generalization – specifically, the scope configuration in the ellipsis PD lacks a parallel configuration in the antecedent PD (recall our discussion of the comparable sequence in (15) above).

We conclude that an analysis of CEPs on which Uniqueness is encoded in the meaning of the CEP cannot correctly capture the observed interpretation of sequence (31). In particular, although one obtains the right interpretation of the second sentence if one assumes that the quantifier containing the CEP scopes above the modal, this representation runs afoul of the independently supported generalizations about the behavior of scope-bearing elements in ellipsis contexts, that is, Scope Parallelism Generalization.

Trapping Uniqueness. We saw that the Uniqueness inference may be absent from the ellipsis PD, even if there is a CEP in the antecedent PD. But it, of course, need not be. That is, an exceptive anteceded by a CEP may trigger Uniqueness. And, moreover, in some cases it has to trigger it. Such a case is exemplified by the sequence in (35), where the first sentence should be parsed as conveying that I performed imperfectly in every test (= where *every exercise but the last one* takes scope below *every test*).

(35) I solved every exercise but the last one in every test.
?To get an A for the course, you really had to △.

The second sentence in this sequence appears to only be able to convey that my imperfect performance was required. This reading is pragmatically marked in natural contexts in which the more exercises you solve the better you do in a test and course. This is straightforwardly predicted by von Fintel's (1993) account, as indicated in (36)-(37): the CEP is contained in the antecedent PD, and it must thus be generated in a parallel position at the ellipsis PD; there it gives rise to the odd meaning described above.

(36) a. [every test] [λt [[every excercise but the last one] [λx [I solved x in t]]]]
b. [$\Box_F$ [[every test] [λs [[every ex. but the last one] [λy [you$_F$ solved y in s]]]]]]

2.4 NP ellipsis

Finally, consider sequence (37), in which an NP in the second sentence is elided.

(37) While Mary aced every course but her electives, most boys only aced a few △. However, every boy did ace almost all of his electives.

The second sentence in (37) conveys the same meaning as the sentence in (38). In particular, the sentence can be judged as true even if every boy aced (almost) all of their electives, as witnessed by the felicity of the final sentence. This is the case also in scenarios in which the majority of the boys' classes were electives.

(38) Most boys only aced a few courses other than their electives.

Puzzle. This state of affairs is unexpected on the theory of CEPs introduced above. The first sentence in (37) has the form provided in (39), where the subject binds the pronoun contained in the CEP.

(39) [Mary [λx [every course but her$_x$ electives] [λy [x aced y]]]]

Three representative candidates for the LF of the second sentence are provided in (40): the first representation contains a CEP and the ellipsis PD is taken to be the matrix clause (a proper subconstituent of the λz-prefixed constituent could not satisfy Parallelism due to No Meaningless Coindexation); the second representation lacks an exceptive and the ellipsis PD is taken to be the NP alone; and the third representation contains an *other than* exceptive and, again, the PD is taken to be the matrix clause. None of these representations forms a licit sequence with the sentence in (39) however, which is why the felicity of (37) is unexpected. (We make some simplifying assumptions about the structure and focus marking of the DP containing the elided NP.)

(40) a. [most b.]$_F$ [λz [only$_F$ [a few$_F$ courses but their$_z$ electives] [λu [z aced u]]]]
b. [most b.]$_F$ [λz [only$_F$ [a few$_F$ courses] [λu [z aced u]]]]
c. [most b.]$_F$ [λz [only$_F$ [a few$_F$ courses other than their$_z$ el.] [λu [z aced u]]]]

First: the exceptive in (40a) results in a contradictory meaning, given that a CEP modifies an existential quantifier. Second: although the PD in (40b) trivially satisfies Parallelism if the antecedent PD is chosen to be the NP *courses*, it does not convey the meaning that we observe – namely, one does not quantify over all the courses but rather only the courses that are not the respective boy's electives (again, as witnessed by the final sentence in the sequence).[6] Third: the structure in (40c) has a consistent meaning, but does not satisfy Parallelism, as indicated in (41). Specifically, none of the focus alternatives to the structure in (40c) induces Uniqueness, which we observe for sentence (39).

(41) *Parallelism Condition is violated*:
⟦[Mary [λx [every course but her$_x$ electives] [λy [x aced y]]]]⟧ ∉ ALT([most boys]$_F$ [λz [only$_F$ [a few$_F$ courses other than their$_z$ electives] [λu [z aced u]]]])

[6]One could try to obtain the correct meaning by relying on the covert resource domain argument of the indefinite *a few* (von Fintel 1994, Stanley & Szabó 2000, among others). To achieve this, one would need to assume that the resource domain argument is effectively bound by the subject, as represented in (i), where $f_D(x)$ would pick out a set of individuals that does not contain x's electives. Although this path is indeed possible, it faces a variety of issues carefully documented by Elbourne 2005, 2013, and others. We will see that the account we outline in Section 4 skirts these issues.

(i) [most boys]$_F$ [λy [only$_F$ [a few$_F$ [[f$_D$ pro$_y$] courses]] [λz [y aced z]]]]

We conclude that an analysis of CEPs on which Uniqueness is encoded in the meaning of the CEP cannot correctly capture the observed interpretation of the sequence in (37): every construal of the sentences either violates Parallelism or derives meanings that are different from the desired ones.

2.5 Diagnosis

We presented three arguments against analyses of CEPs that take the mechanism governing their distribution to be encoded in the meaning of the CEPs themselves, in particular, an analysis that encodes both Uniqueness and Domain Subtraction into the CEPs.

In all the examples that we discussed, the predicament turned out to be that one can (i) satisfy the conditions on ellipsis licensing – by having Uniqueness triggered in the alternatives – but derive an undesired interpretation – by triggering Uniqueness at the ellipsis site; or (ii) derive the desired interpretation – by not triggering Uniqueness at the ellipsis site – but fail to satisfy the conditions on ellipsis licensing – by not triggering Uniqueness in the alternatives. A schematic of the structures that satisfy Parallelism (but fail to have the desired interpretation) is provided in (42).

(42) a. [$_S$ ... [$_{\text{Antecedent PD}}$ ... [$_{QP}$ every/no/some NP but NP'] ...] ...]

b. [$_S$ ... [(OP) ... [$_{\text{Ellipsis PD}}$ ... [$_{QP}$ every/no/some NP but NP'] ...] ...] ...]

↓

Subtraction+Uniqueness

However, if CEPs picked out solely Domain Subtraction, and the mechanism governing their distribution, corresponding to Uniqueness, were generated outside the antecedent PD, the observed predicament would disappear: since the regulative inferences accompanying CEPs would not be generated in the antecedent PD, they would not need to be generated in the ellipsis PD either. This would effectively mean that a CEP could in principle antecede an *other than* exceptive that does not exhibit Uniqueness.

3 *Almost*

In this section we show that a similar puzzle arises for approximative modifier *almost*. Approximative modifiers give rise to two inferences: Proximal and Negative Implication, as exemplified in (43). And they are subject to a similar constraint on distribution that we observed for CEPs: they are acceptable when they modify universal quantifiers, but not when they modify existential ones, as exemplified in (44).

(43) a. Almost every book is worth reading.
b. *Proximal Implication*: Close to every book is worth reading.
c. *Negative Implication*: Not every book is worth reading.

(44) a. Almost every book is worth reading.
b. Almost no book is worth reading.
c. *Almost some book is worth reading.

Some approaches to *almost*. A prominent class of approaches to *almost* takes the semantic contribution of approximatives to a sentence – Negative and Positive Implication – to be encoded in the meaning of the approximatives themselves (e.g., Morzycki

2001, Penka 2006, Nouwen 2006, among others). For example, the meaning that these approaches assign to *almost* is something along the lines of (45): in the first characterization, *almost* modifies the quantificational determiner and conveys that a quantifier close to it holds of its arguments (= Proximal Implication), but the quantifier itself does not (= Negative Implication) (cf. Morzycki 2001; see also discussion in von Fintel 1994, Moltmann 1995, Keenan 1996); in the second characterization, *almost* modifies the clause and conveys that a proposition close to it is true (= Proximal Implication), but the clause itself is false (= Negative Implication) (cf. Penka 2006). (We abstract away from some details of the specific approaches. We assume that the alternatives on which clausal *almost* operates on are exclusively those induced by the quantificational determiner.)

(45) a. $[\![\text{almost}]\!](Q)(P)(R) = 1$ iff $\exists Q'$ (Q' is close to Q $\wedge$ Q'(P)(R)) $\wedge \neg Q(P)(R)$
b. $[\![\text{almost}]\!](p) = 1$ iff $\exists q(q \text{ is close to } p \wedge q) \wedge \neg p$

These meanings allow one to capture Proximal and Negative Implications of sentences with *almost*, as well as derive their acceptability only with certain types of quantifiers. For illustration, on the one hand, the meanings contributed by *almost* in *almost every* and *almost no* are contingent, as shown in (46)-(47) (say, the sentences are true when every book but one is worth reading and when only one book is worth reading, respectively).

(46) a. Almost every book is worth reading.
b. $\underbrace{\text{ˆ(Close to every book is worth reading)}}_{\text{Proximal Implication}} \wedge \underbrace{\neg\text{(Every book is worth reading)}}_{\text{Negative Implication}}$

(47) a. Almost no book is worth reading.
b. ˆ(Close to no book is worth reading) $\wedge \neg$(No book is worth reading)

On the other hand, the meaning contributed by *almost* in *almost some* in (48) is inconsistent, at least on the assumption that the alternatives on which *almost* operates must be compatible with the (quantificational or propositional) argument of *almost* (see, e.g., Penka 2006, Nouwen 2006, and others).

(48) a. *Almost some book is worth reading.
b. ˆ(Close to some book is worth reading) $\wedge \neg$(Some book is worth reading)

The above proposal is thus able to account both for the semantic contribution of *almost* and its distribution in a compositional way. In the following, however, we argue on the basis of ellipsis data that whereas Proximal Implication of *almost* is indeed induced by *almost* in the DP (or the VP) in which it occurs, Negative Implication must be taken to be triggered outside the VP. (See Sadock 1981 for an early view along this line.)

Puzzle. Consider (49).

(49) In the exam, I solved almost every exercise. To get an A, you really had to $\triangle$.

The sequence is felicitous in natural contexts in which the more exercises you solve, the better you do in an exam. More to the point: the first sentence in the sequence triggers both Proximal and Negative Implication, as given in (50); the second sentence, on the other hand, entails merely (51), that is, it does not have to trigger Negative Implication in the scope of the modal.

(50) a. I solved close to every exercise.
b. I did not solve every exercise.

(51) a. You had to solve (at least) close to every exercise.
b. You did not have to solve every exercise.

The structure of the first sentence in (49) is one of (52a)-(52b), depending on whether *almost* is taken to modify a quantificational determiner or a clause. The interpretation of the two structures is computed in (53), which corresponds to (50).

(52) a. [[almost every] exercise] λx [I solved x]]
b. [almost [[every exercise] λx [I solved x]]]

(53) $\hat{}$(I solved close to every exercise) $\wedge$ $\neg$(I solved every exercise)

In order to satisfy Parallelism, the matrix clause in (52) must be taken to be the antecedent PD and the sentential complement of the second sentence to be the ellipsis PD, as given in (54), where, again, *almost* modifies a quantificational determiner or a clause. Recall that if either *almost every exercise* or just *almost* take scope above the modal, we obtain a violation of Scope Parallelism Generalization.

(54) a. [$\Box_F$ [[almost every] exercise] λy [you$_F$ solved y]]]
b. [$\Box_F$ [almost [[every exercise] λy [you$_F$ solved y]]]]

The meaning of the structures in (54), given in (55), does not correspond to the meaning of the second sentence in (49): namely, the structures in (54) entail that you were required not to solve every exercise to get an A.

(55) $\Box\big(\hat{}$(I solved close to every exercise) $\wedge$ $\neg$(I solved every exercise)$\big)$

Summary and diagnosis. An approach to *almost* that assumes that that both Proximal and Negative Implication are encoded in the meaning of *almost* faces the same issues that an approach that encodes both Domain Subtraction and Uniqueness in the meaning of CEPs does, namely, it undergenerates when it comes to ellipsis contexts. Specifically, on the common assumptions about ellipsis, such an approach incorrectly predicts that an approximative modifier anteceded by *almost* will trigger Negative Implication at the ellipsis site, in particular, that it will trigger Negative Implication in a scope position that is structurally parallel to that of *almost* in the antecedent PD, as schematized in (56).

(56) a. [$_S$ ... [$_{\text{Antecedent PD}}$ ... [$_{VP}$... almost ...] ...] ...]
b. [$_S$ ... [(OP) ... [$_{\text{Ellipsis PD}}$ ... [$_{VP}$... almost ...] ...] ...] ...]
$\downarrow$
Proximal+Negative

Similar to what we observed for CEPs, if *almost* picked out solely Proximal Implication and the mechanism governing its distribution, corresponding to Negative Implication, were generated outside the antecedent PD, we would obtain the desired meanings: since the regulative inferences accompanying *almost* would not be generated in the antecedent PD, they would not need to be generated in the ellipsis PD either. This would effectively mean that *almost* could in principle antecede an approximative that does not induce Negative Implication.

4 Separation

We showed that the distribution of CEPs and *almost* in ellipsis contexts is incompatible with these elements encoding Uniqueness and Negative Implication, respectively. Namely, on this assumption, we predict that these inferences are necessarily generated both in the antecedent constituent and at the ellipsis site, in contrast to what we observe.

A desired resolution of the sequences under discussion may, however, be obtained if Uniqueness and Negative Implication were not induced in the PDs, but rather higher in the structure. This turns out to be compatible with, and thus provide support for, Gajewski's (2013) approach to CEPs (but not Gajewski's 2008 approach) and Spector's (2014) approach to *almost*. On these approaches, CEPs and *almost* have a simple monotone semantics, and the inferences corresponding to Uniqueness and Negative Implication, which govern their distribution, are induced by an exhaustivity operator that may be generated higher in the structure, safely outside potential PDs. We briefly appraise these two approaches in light of the ellipsis puzzles.

4.1 Connected exceptives

Gajewski (2008, 2013) argues for an approach to CEPs on which Domain Subtraction is triggered by the CEP, while the inference corresponding to Uniqueness is generated elsewhere in the structure. He motivates this split by looking at two phenomena: (i) felicitous modification of NPIs by CEPs, and (ii) licensing of NPIs in the scope of negative quantifiers modified by CEPs. Instead of reproducing Gajewski's ingenious treatment of these phenomena, we focus on how his approach fares with the ellipsis puzzles. (We modify aspects of Gajewski's analysis in order to simplify our presentation. The reader is referred to his papers cited above for full details.)

Domain Subtraction and Exhaustivity. Gajewski (2013) assigns a CEP like *but 'War and Peace'* the function of subtracting 'War and Peace' from the denotation of its sister at LF, as given in (57). Exceptives like *other than NPs* function the same way.

(57) $[\![\text{but 'War and Peace'}]\!] = [\![\text{other than 'War and Peace'}]\!] = \lambda P.\ P \backslash \{WP\}$

Furthermore, Gajewski assumes that CEPs are obligatorily associates of covert exhaustivity operator *exh* (see also Gajewski 2008, Sect. 3.2.2, and Hirsch 2016). The meaning of *exh*, which has been independently argued to be responsible for generating scalar implicatures and to be involved in NPI licensing, is provided in (58) (see Fox 2007 for details, and Krifka 1995, Chierchia 2013 for its role in NPI licensing).

(58) $\text{exh}(S) = 1$ iff ${}^{\vee}[\![S]\!] \wedge \forall q \in \text{ALT}(S)\ [\ {}^{\vee}q \rightarrow ([\![S]\!] \Rightarrow q)\ ]$

Sample derivation. The import of *exh* consists of negating the non-weaker alternatives that are induced by the CEP. These alternatives are identical to the exception sentence except for the individual picked out by the exceptive being substituted by other individuals. For example, *Every book but 'War and Peace' is worth reading* has the proposition that every book that is not 'Anna Karenina' is worth reading, etc., as an alternative, as given in (59) (see Hirsch 2016 for further discussion).

(59) ALT(Every book but 'War and Peace' is worth reading) =
$\{{}^{\wedge}(\text{every book that is not X is worth reading}) \mid X \in D_e\}$

Accordingly, the resulting exhaustification effectively involves universal quantification over all the individuals distinct from the excepted element. The sentence in (60) thus has the interpretation in (61). (Gajewski 2013 and Hirsch 2016 show in detail how one gets contingent meanings when and only when the CEPs are judged to be grammatical.)

(60) a. Every book but 'War and Peace' is worth reading.
b. [exh [[every book but 'War and Peace'] [is worth reading]]]
association with alternatives

(61) $\hat{}$(Every book that is not WP is worth reading) $\wedge$ [= Subtraction]
$\forall$X $\neq$ WP: $\neg$(Every book that is not X is worth reading) [= Exhaustivity]

The meaning in (61) entails that 'War and Peace' is a book – otherwise the meaning would be contradictory since it would convey that every book is worth reading but that not every book that is not, say, 'Anna Karenina' is worth reading – and that 'War and Peace' is not worth reading – if it were worth reading, it could not hold that the prejacent of *exh* is true and that, say, not every book that is not 'Anna Karenina' is worth reading. These inferences correspond to those derived in von Fintel's (1993) proposal.

(62) *Inferences following from (61):*
a. Every book that is not 'War and Peace' is worth reading.
b. 'War and Peace' is a book.
c. 'War and Peace' is not worth reading.

Ellipsis puzzles. Covert exhaustification has one characteristic that is crucial for the derivation of the ellipsis puzzles (see Fox 2004): when it applies to alternatives induced in an antecedent PD, it does not have to be contained in the PD itself (this holds not least because it does not stand in a binding/movement dependency relation to its associate). This explains why the first sentence in (63) may be exhaustified, while the second sentence need not be, esp. if there is a 'suspender' continuation. The pertinent LFs are in (64).

(63) John solved some of the exercises. Bill did $\triangle$ too. (In fact, he solved all of them.)

(64) a. [exh [John solved some of the exercises]]
b. [Bill$_{\text{F}}$ solved some of the exercises]

Turning now to the ellipsis puzzles, one can show that all the data fall out naturally from the above assumptions. We focus solely on the case of universal quantifiers in the following. Recall sequence (31), repeated below, in which the exceptive modifying the universal quantifier in the elided VP does not induce the standard inferences accompanying the CEPs, at least not in the scope of the modal.

(31) In the exam, I solved every exercise but the last one.
To get an A, you really had to $\triangle$.

The first sentence in (31) is assigned the representation in (65). The sister of *exh* can be chosen as the antecedent PD. The sentence entails that I solved every exercise that is not the last one and that I did not solve the last exercise, along the lines sketched above.

(65) a. I solved every exercise but the last one.
b. [exh [[every exercise but the last one] [λx [I solved x]]]]

The second sentence in the sequence may be assigned the representation in (66). The sentential complement of the modal can be chosen as the ellipsis PD. The exhaustification of the *other than* exceptive, which is optional, as indicated by the parantheses (66), can be induced at the matrix level, outside the ellipsis PD. Obviously, Parallelism is satisfied by these representations, as indicated in (67).

(66) a. To get an A, you really had to △.
b. [(exh) [$\Box_F$ [[every exercise other than the last one] [λy [you$_F$ solved y]]]]]

(67) *Paralellism Condition is satisfied*:
⟦[every exercise but the last one] [λx [I solved x]]]⟧ ∈
ALT([every exercise other than the last one] [λy [you$_F$ solved y]])

If the exhaustification applies in (66b), we obtain the meaning that you had to solve every exercise that is not the last one to get an A, and that you did not have to solve the last exercise, which is equivalent to the meaning computed in (68). This indeed corresponds to the intuitive meaning of the sentence and coheres with the natural assumption that the more exercises you solve, the better you do in an exam.

(68) $\hat{}\Box$(you solve every exercise that is not L) ∧ [= Subtraction]
∀X ≠ L: ¬□(you solve every exercise that is not X) [= Exhaustivity]

Trapping Uniqueness. We discussed sequence (35), repeated below, where the second sentence is pragmatically marked since it requires imperfect performance on every test.

(35) I solved every exercise but the last one in every test.
?To get an A for the course, you really had to △.

We have seen in Sect. 2.3 that the availability of solely this reading follows naturally on von Fintel's approach. However, it is also predicted on Gajewski's approach: namely, on the desired resolution, the first sentence in (68) is parsed as in (69a), where *exh* is generated in the scope of *every test*. Accordingly, *exh* is contained the antecedent PD and thus the ellipsis PD as well.

(69) a. [every test] [λt [exh [[every exercise but L] λx [I solved x in t]]]]
b. [$\Box_F$ [[every test] [λt' [exh [[every exercise but L] λx' [you$_F$ solved x' in t']]]]]]

The meaning of the structure in (69b) is pragmatically marked: it expresses a requirement that every test is such that I did not solve the last exercise on it (in addition to solving every other exercise on it). Thus, we see that *exh* may not only be optionally generated at the ellipsis site,[7] in some cases it must be present to satisfy Parallelism.

[7] In simple sequences like (i), there is a strong preference for the second sentence to induce Uniqueness. That is, there is a preference to exhaustify the alternatives induced by the elided exceptive, as given in (ii). This is unsurprising in light of an independent preference – not a requirement! – to exhaustify the alternatives induced by *other than* exceptives.

(i) John liked every book but 'War and Peace'. Mary did △ too.

(ii) a. [exh [[every book but 'War and Peace'] [λx [John liked x]]]]
b. [exh [[every book other than/but 'War and Peace'] [λx [Mary$_F$ liked x]]]]

Gajewski 2008. There is one crucial difference between Gajewski's (2008) approach to CEPs and his 2013 proposal: the former is implemented in a way that requires a movement/binding dependency relation to obtain between the CEP and the mechanism governing its distribution. For illustration, on Gajewski's (2008) approach, sentence (70) has the representation in (71), where Uniqueness-inducing operator *Unique* takes matrix scope and the excepted phrase moves to its complement position at LF (another option would be to move to its specifier position, see Wagner 2006 for discussion).

(70) I solved every exercise but the last one.

(71) [Unique L] [λx [[every exercise but x] [λy [I solved y]]]]

The meaning of the structure is identical to what we computed above (*Unique* is engineered so as to yield meanings identical to those of von Fintel's proposal for unembedded sentences). Importantly for us, the representation in (71) forces us to select the matrix clause as the antecedent PD in the sequence in (31), repeated below.

(31) In the exam, I solved every exercise but the last one.
To get an A, you really had to △.

The structure of the second sentence must then be one of the following:

(72) a. [$\Box_F$ [[Unique L] [λz [[every exercise but z] [λu [you$_F$ solved u]]]]]]
b. [Unique L] [λz [$\Box_F$ [[every exercise but z] [λu [you$_F$ solved u]]]]]

In parallel to our discussion in the preceding section, while such structures may satisfy Parallelism, they either fail to convey the desired interpretation – this is the case for (72a), which conveys that you were required not to solve the last exercise – or they violate Scope Parallelism Generalization – this is the case for (72b), in which the *Unique*-phrase takes scope above a universal modal in the ellipsis PD without there being a parallel scope configuration in the antecedent PD.

More generally, when it comes to CEPs in ellipsis contexts, any approach that assumes a movement (or binding) dependency relation between the mechanism governing the distribution of CEPs and the CEP is in the same boat as the approaches that encode this mechanism into the CEP. The main problem for such approaches lies in the fact that binding into (/movement out of) the antecedent constituent requires parallel binding into (/movement out of) the elided constituent (cf. Scope Parallelism Generalization).

Summary. We showed that the ellipsis puzzles are naturally resolved on Gajewski's (2013) approach to CEPs, on which Uniqueness is peeled off from the CEP and its function is taken over by *exh*. On this approach, a sentence containing the antecedent PD may be parsed so that *exh* accompanying the CEP is not contained in the PD (unless one is dealing with a trapping configuration). This means that *exh* need not be contained in the ellipsis PD, though it may be so (and in the trapping configurations must be so). Finally, we argued that a derivation that relies on a movement/binding dependency

Some of the questions that arise in light of such patterns are (i) why exhaustification with *other than* exceptives is only optional (though preferred) and (ii) why exhaustification with CEPs is obligatory. Another difference between the two types of exceptives that we would like to understand relates to how local *exh* has to be to the exceptive (when it applies). We cannot pursue these issues here. See Gajewski 2013, Sect. 4, and Gajewski 2008, Sect. 3.4, for some discussion.

relation between the mechanism governing its distribution and the CEP is not able to adequately resolve the ellipsis puzzles. Rather, the relation between the two must be of the kind put forward by some authors for association with focus and alternatives (see esp. Rooth 1985, 1992, Fox 2007).

4.2 *Almost*

Spector (2014) proposed to analyze approximative modifiers like *almost* as encoding only Proximal Implication, and to derive Negative Implications associated with them via *exh*. Spector's proposal can be naturally extended to cover the examples discussed here.

Domain Subtraction and Exhaustivity. In the examples discussed above, we can take *almost* to modify a quantifier and trigger Proximal Implication, as in (73). We assume that a quantification over a set is 'close to' the same quantification over a slightly larger superset (that is, $Q(P \backslash D_c)$ is close to $Q(P)$ on an appropriate selection of D_c).

(73) $[\![\text{almost}]\!](Q)(P) = Q(P \backslash D_c)$, where D_c is small and determined by the context

Negative Implication is triggered by *exh* that c-commands *almost* and uses up the alternatives induced by the *almost*-phrase. For simplicity, we will assume that the only relevant alternative is in our cases the quantifier with an intact domain. This means that, for example, *almost every exercise* effectively has *every exercise* as the alternative.

Sample derivation. The sentence in (74a) is assigned the structure in (74b), where *exh* takes matrix scope and associates with the alternatives induced by *almost*.

(74) a. I solved almost every exercise.
b. [exh [[[almost every] exercise] [λx [I solved x]]]]

association with alternatives

The meaning of the structure in (74b) is provided in (75): it conveys that I solved close to every exercise, which corresponds to the meaning of the sister of *exh*, and that I did not solve every exercise, which corresponds to the negation of the alternative induced by *almost*. The latter inference is triggered by *exh*.[8]

[8]The assumptions introduced above suffice to derive the restricted distribution of *almost* in quantifiers. In particular, when *almost* modifies *some*, the resulting meaning is contradictory. This is shown in (i):

(i) a. *Almost some book is worth reading.
b. [exh [[[almost some] book] is worth reading]]
c. ˆ(some book that is not in D is worth reading) ∧ ¬(some book is worth reading) [= ⊥]

Now, *almost* has a broader syntactic distribution than CEPs, for example, it can modify VPs. These other occurrences of *almost* arguably have a semantics closely related to that provided in the main text. For example, Spector (2014, Sect. 6) assigns *almost* the meaning along the lines of (ii), where P is a predicate of degrees, and d is a degree argument of P (say, the maximal degree of which P can hold); the function of *almost* is to subtract some small context-dependent value from the degree argument.

(ii) $[\![\text{almost}]\!](P)(d) = 1$ iff P(d-d'), where d' is small and determined by the context

The predicate applying to the subtracted-from degree, P(d), is an alternative to the *almost*-phrase that is logically stronger than it (namely, P(d) entails P(d-d')), and is subsequently negated by *exh*.

(75) ˆ(I solved every exercise that is not in D) ∧ [= Proximal Implication]
¬(I solved every exercise), [= Exhaustivity]
where D is small and determined by the context

Ellipsis puzzle. The sequence in (49), repeated below, is pragmatically plausible. The second sentence in the sequence conveys that you were required to solve close to every exercise, but were not required to solve every exercise in order to get an A.

(49) In the exam, I solved almost every exercise. To get an A, you really had to △.

This interpretation can be derived from the structures in (76). The antecedent PD consists of the sister of *exh*, while the ellipsis PD consists of the sentential complement of the modal. This configuration trivially satisfies Parallelism, as indicated in (77).[9]

(76) a. [exh [[[almost every] exercise] [λx [I solved x]]]]
b. [exh [$\Box_F$ [[[almost every] exercise] [λy [you$_F$ solved y]]]]]

(77) *Parallelism Condition is satisfied*:
⟦[[[almost every] exercise] [λx [I solved x]]]⟧ ∈
ALT([[[almost every] exercise] [λy [you$_F$ solved y]]])

Moreover, the structure of the second sentence, in which *exh* takes matrix scope, conveys precisely the meaning described above, given in (78).

(78) ˆ□(you solved every exercise that is not in D) ∧ ¬□(you solved every exercise),
where D is small and determined by the context

Summary. We showed that the ellipsis puzzle with *almost* does not arise on an approach on which Negative Implication accompanying *almost* is not encoded in the meaning of *almost*, but is rather triggered by an operator higher in the structure. This is in line with Spector's (2014) proposal, who takes this operator to be *exh*.

5 Conclusion

We developed an argument based on ellipsis against encoding the mechanisms governing the distribution of CEPs and *almost* into their lexical meaning. Rather, their distribution must be taken to be governed by semantic means external to them. Moreover, we showed that these means may not stand in a movement (or binding) dependency relation with them. Finally, we pointed to two proposals compatible with our conclusions: Gajewski's (2013) approach to CEPs and Spector's (2014) approach to *almost*.

The analysis of the ellipsis data involving CEPs and *almost* that we argued for in this squib thus resembles the standard treatments of other types of alternations in ellipsis, in particular, alternations involving NPIs (e.g., Sag 1976, Merchant 2013). As with CEPs, other examples of alternations tend to be analyzed as involving an element at the ellipsis

[9] Unlike in the ellipsis examples with exceptives, we assume that the same approximative expression, *almost*, is used both in the antecedent and the ellipsis PD. The main reason for using *other than NP* in the ellipsis PDs with exceptives was that it – in contrast to *but NP* – does not necessitate a local *exh* operator. We make a substantive assumption in this section that *almost* does not either. We hope to examine this assumption in greater detail in the future. See Spector 2014 for some preliminary support.

site that has a meaning identical to that of their antecedent but whose distribution (or morphological realization) is conditioned by an operator higher in the structure.

References

Abels, Klaus & Luisa Martí. 2010. A unified approach to split scope. *Natural Language Semantics* 18(4). 435–470.

Beck, Sigrid. 2012. Degp scope revisited. *Natural language semantics* 20(3). 227–272.

Chierchia, Gennaro. 2013. *Logic in grammar*. Oxford: Oxford University Press.

van Craenenbroeck, Jeroen & Jason Merchant. 2013. Ellipsis phenomena. In Marcel den Dikken (ed.), *The cambridge handbook of generative syntax*, 701—745. Cambridge University Press.

van Craenenbroeck, Jeroen & Tanja Temmerman. 2015. How (not) to elide negation. Manuscript, KU Leuven.

Crnič, Luka. 2016. Free choice under ellipsis. Manuscript, The Hebrew University of Jerusalem.

Elbourne, Paul. 2005. *Situations and individuals*. Oxford University Press.

Elbourne, Paul. 2013. *Definite descriptions*. Oxford: Oxford University Press.

Fiengo, Robert & Robert May. 1994. *Indices and identity*, vol. 24. MIT press.

von Fintel, K. 1994. *Restrictions on quantifier domains*: University of Massachusetts dissertation.

von Fintel, Kai. 1993. Exceptive constructions. *Natural Language Semantics* 1(2). 123–148.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. MIT Press.

Fox, Danny. 2004. Lectures on scalar implicatures. Lecture notes, MIT.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 71–120. Palgrave Macmillan.

Gajewski, Jon. 2002. L-analyticity and natural language. Manuscript, MIT.

Gajewski, Jon. 2008. Npi any and connected exceptive phrases. *Natural Language Semantics* 16(1). 69–110.

Gajewski, Jon. 2013. An analogy between a connected exceptive phrase and polarity items. *Beyond'Any'and'Ever': New Explorations in Negative Polarity Sensitivity* 262. 183.

Griffiths, James & Anikó Lipták. 2014. Contrast and island sensitivity in clausal ellipsis. *Syntax* 17(3). 189–234.

Hartman, Jeremy. 2011. The semantic uniformity of traces: Evidence from ellipsis parallelism. *Linguistic Inquiry* 42(3). 367–388.

Heim, Irene. 1996. Lectures on ellipsis. Lecture note, MIT.

Hirsch, Aron. 2016. An unexceptional semantics for expressions of exception. In *University of pennsylvania working papers in linguistics*, vol. 22 1, .

Johnson, Kyle. 2001. What vp-ellipsis can do, and what it can't, but not why. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 439–479. Oxford: Blackwell.

Keenan, Edward L. 1996. The semantics of determiners. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 41–64. Malden, MA: Blackwell.

Krifka, Manfred. 1995. The semantics and pragmatics of weak and strong polarity items. *Linguistic Analysis* 25. 209–257.

Merchant, Jason. 2013. Polarity items under ellipsis. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosing syntax*, 441–462. Oxford: Oxford University Press.

Moltmann, Friederike. 1995. Exception sentences and polyadic quantification. *Linguistics and Philosophy* 18(3). 223–280.

Morzycki, Marcin. 2001. " almost" and its kin, across categories. In *Semantics and linguistic theory*, vol. 11, 306–325.

Nouwen, Rick. 2006. Remarks on the polar orientation of almost. In *Linguistics in the netherlands*, vol. 23 1, 162–173. John Benjamins Publishing Company.

Penka, Doris. 2006. Almost there: The meaning of almost. In *Proceedings of sinn und bedeutung*, vol. 10, 275–286.

Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Rooth, Mats. 1985. *Association with focus*: University of Massachusetts, Amherst, PhD dissertation.

Sadock, Jerrold M. 1981. Almost. *Radical pragmatics* 257–271.

Sag, Ivan. 1976. *Deletion and logical form*: Massachusetts Institute of Technology, PhD dissertation.

Schwarzschild, Roger. 1999. Givenness, AVOIDf and other constraints on the placement of accent. *Natural Language Semantics* 7. 141–177.

Spector, Benjamin. 2014. Global positive polarity items and obligatory exhaustivity. *Semantics and Pragmatics* 7(11). 1–61.

Stanley, Jason & Zoltan Gendler Szabó. 2000. On quantifier domain restriction. *Mind & Language* 15(2-3). 219–261.

Wagner, Michael. 2006. Association by movement: evidence from npi-licensing. *Natural Language Semantics* 14(4). 297–324.