Exceptive-additive *besides*: a unified semantic treatment*

Clemens Mayr
University of Göttingen

Ekaterina Vostrikova
University of Götingen

November 3, 2026

Abstract

This paper deals with the exceptive-additive expression *besides*. Where the better-studied pure exceptive expression *but* contributes an inference of exception (e.g., with *every* or *no*), *besides* does so too. In environments where *but* leads to ungrammaticality, however, *besides* is acceptable and systematically contributes an inference of addition (e.g. with *some*, *at most three*, *who*). A unified account explaining this behavior of *besides* and its differences from *but* is proposed. The account extends the treatment of *but* in terms of *EXH* (Gajewski 2013, Hirsch 2016, Crnič 2018, 2021, 2024): while *besides* and *but* both contribute domain subtraction, the resulting quantificational claims get strengthened differently in the two cases by *EXH*. The source of this lies in a difference in how the alternatives for *EXH* are generated. This—coupled with application of *EXH* and a maximization operation in some instances—explains the patterns found for *besides* and *but*.

1 Introduction

Exceptive constructions, such as those introduced by *except* and *but* in English, have received significant attention and discussion in the literature. In this paper, our focus lies on the semantics of a less explored construction introduced by *besides*. While it shares some core properties with exceptives, it also differs substantially from them.

The two main semantic properties of exceptive constructions are their uniform applicability to both universal and negative quantifiers, contributing an exceptive inference in such contexts (Hoeksema 1987, von Fintel 1993), and their inability to combine with quantifiers of any other quantificational force (Horn 1989). *Besides* behaves like exceptives concerning the first property, but crucially, does not possess similar restrictions on quantifier compatibility.

*We would like to thank Luka Crnič, Kai von Fintel, Nina Haslinger, Aron Hirsch, Uli Sauerland, Benjamin Spector, Winfried Lechner, Hedde Zeijlstra and the audiences at Göttingen University, the University of Düsseldorf, the University of Tübingen, the workshop on "Alternatives, Expectations, and Domain Widening" at Bar-Ilan University, Creteling 2022 and Sinn und Bedeutung 27 for comments, suggestions, and criticism. The research reported here was funded by DFG (Deutsche Forschungsgemeinschaft) grant STE 2555/4-1 awarded to Clemens Mayr. Vostrikova's contribution was funded by DFG (grant STE 2555/4-1) and by the European Union (ERC grant UNPAG 101142366). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

To illustrate the first property of pure exceptives, let's consider examples (1) and (2). In (1), where the quantificational generalization is positive, the use of *but* or *except* results in a negative exception inference. Conversely, with the negative quantificational generalization in (2), the exception inference becomes positive. Additionally, both *but* and *except* contribute the inference that the individual to which the exceptive inference applies belongs to the restrictor set. This particular inference is referred to as *containment* in the exceptive literature. In this case, this is the inference that Ann is a student.

(1) Every student *but/except/besides* Ann passed.
 ⇝ Every student who is not Ann passed — quantification
 ⇝ Ann didn't pass — exception
 ⇝ Ann is a student — containment

(2) No student *but/except/besides* Ann passed.
 ⇝ No student who is not Ann passed — quantification
 ⇝ Ann passed — exception
 ⇝ Ann is a student — containment

As is also seen from (2), *besides* used in these environments behaves identically. Not so, however, when it comes to the second key semantic property of exceptive constructions: as illustrated in (3) pure exceptives are unacceptable with non-universal quantifiers (Horn 1989, von Fintel 1993).[1]

(3) a. *Some student *but/except* Ann passed.
 b. *At least one/exactly one student *but/except* Ann passed.
 c. *At most one/more than one student *but/except* Ann passed.

Besides is grammatical in such cases, as shown in (4), (5), and (6). Importantly, the inference contributed by *besides* in all these cases is positive (Ann passed). In this regard, these cases are similar to the case with *no student* in (2), even though no negative quantifier is present in (4) and (5). The sentence in (4) can be paraphrased as follows: in addition to Ann, some other students passed. We will call these readings additive and categorize *besides* as an exceptive-additive expression, in contrast to pure exceptives like *but* and *except*. The containment inference is present in all cases. In this regard the exceptive and the additive uses of *besides* pattern like exceptives.

(4) Some student(s) *besides* Ann passed.
 ⇝ Some student(s) who is not Ann passed — quantification
 ⇝ Ann passed — addition
 ⇝ Ann is a student — containment

(5) At least/more than/exactly one student *besides* Ann passed.
 ⇝ At least/more than one student who is not Ann passed — quantification
 ⇝ Ann passed — addition
 ⇝ Ann is a student — containment

[1] There is debate in the literature regarding the compatibility of exceptive constructions with quantifiers that are close to the endpoint of a Horn scale, such as *many*, *most*, and *few*. Such expressions might form an exception to this generalization (compare the positions of von Fintel 1994, Moltmann 1995, Gajewski 2013, Hirsch 2016 with those of García Álvarez 2008, Peters and Westerståhl 2006, Nadathur and Lassiter 2018, Peters and Westerståhl 2023, also cf. von Fintel 1994 on *except for* in this regard).

(6) At most/fewer than two students *besides* Ann passed.
⇝ At most/fewer than two students who is not Ann passed quantification
⇝ Ann passed addition
⇝ Ann is a student containment

Another context in which the behavior of *besides* differs from that of exceptives is in *which*-questions. While *except* and *but* are ungrammatical in *which*-questions, as demonstrated in (7) (von Fintel 1994), *besides* is grammatical in these contexts and, once again, yields the additive inference (Vostrikova 2019b). The containment inference is present here, as in all cases considered above.

(7) a. *Which student(s) *but*/*except* Ann passed?
b. Which students *besides* Ann passed?
⇝ Which student(s) who is not Ann passed? question
⇝ Ann passed addition
⇝ Ann is a student containment

The empirical generalization emerging from the review of this picture is given in (8).

(8) **The exceptive-additive generalization (EAG)**: Exceptive-additive *besides* behaves in parallel to exceptives in contexts where exceptives are acceptable. In contexts where exceptives are ungrammatical, *besides* is grammatical and obtains the additive reading.

Our objective is to provide a semantic theory deriving the EAG. That is, the aim is to accurately capture the exceptive-like behavior of *besides* with universal and negative quantifiers, as well as its additive behavior with existentials, modified numerals, and wh-questions. Additionally, we aim to provide an explanation for the differences between *besides* and pure exceptives with a focus on *but*.

There is consensus in the literature that the explanation of the restrictions on the use of pure exceptives should be semantic. The reason is that the empirical pattern is predictable from the logical properties of the quantifiers involved. Given the first part of the EAG, this explanation should carry over to exceptive-additive *besides*. Adopting such an account wholesale, however, would also have the consequence that *besides* with non-universal quantifiers should come out as ungrammatical. Notice incidentally that the fact that *besides* always gives rise to the additive interpretation in these contexts again suggests that one is dealing with a systematic and predictable pattern.

We propose a unified treatment for both the exceptive and additive uses of *besides* without recourse to any ambiguity. The core idea is that the meaning of a sentence containing *besides* is derived using the same mechanisms employed in the semantics of pure exceptives: set subtraction and *EXH* (von Fintel 1993, Gajewski 2013, Hirsch 2016, Crnič 2018, 2021, 2024). The distinction between these two types of constructions lies in how the alternatives are constructed: *besides* uses a smaller set of alternatives. In this case, *EXH* asserts the quantificational claim with subtraction made by *besides* and negates the same claim without the subtraction. This is shown to make the right empirical predictions for *every*, *no* and *exactly n* out of the box, as it were. For other modified numerals a decompositional approach is adopted with *EXH* targeting a constituent parallel to the one found with *exactly n* in the scope of the modified numeral. This treatment is suggested to be generalizable to existential quantifiers and wh-questions. To the best of our knowledge, this is the first proposal that comprehen-

sively accounts for all these different uses of *besides*.[2]

The discussion is structured as follows: Section 2 provides a formal background on semantic theories of exceptive constructions, serving as our starting point, and reviews existing theories on *besides*. Section 3 presents our proposal for the semantics of *besides* and explains its interaction with various quantificational expressions. Section 4 considers the predictions made by our account for exceptive *but* and critically discusses some of the theoretical choices made. Section 5 addresses challenges and extensions to the proposed account having to do with the lexico-logical properties of certain predicates, which necessitate a slight modification of our assumptions for the decomposition of modified numerals. Section 6 concludes and provides an outlook to future research.

2 Background on the semantics of exceptives

2.1 The unified treatment of *but* with negative and universal quantifiers

The core idea behind the semantic theory of exceptives is that they impose a restriction on the domain of the quantifier by subtracting an element from the restrictor set (Hoeksema 1987, von Fintel 1993 a.o.). In (9), for example, *but* subtracts *Ann* from the set of students, resulting in an *every*-claim made for a smaller domain.[3]

(9) Every student but Ann passed. $\rightsquigarrow$ Ann didn't pass

This alone does not account for the exception inference. A widely accepted approach to this involves negating alternative quantificational claims that differ from the original only in terms of what is subtracted from the domain (von Fintel 1993, 1994). Various implementations of this idea have been proposed. Current incarnations use the *EXH* operator and focus alternatives (Gajewski 2008, Hirsch 2016, Crnič 2018, 2021, 2024), thereby building on mechanisms with independent motivation through their use in other empirical domains.

According to this approach, (9) would receive an LF like the one shown in (10). The presence of *EXH* is mandatory. This requirement can be implemented either via an agreement mechanism between *but* and *EXH*, or by assuming that *but* obligatorily introduces focus alternatives that must be used by some operator.

(10) [EXH_{Alt} [$_{IP}$ [every [student [but Ann_{F}]]] passed]]

In this semantic framework, the exceptive itself only contributes the inferences of domain subtraction and containment. The lexical entry of *but* can be represented as in (11). It is a function that takes an atomic or plural individual (provided by the DP that follows *but*) and a predicate of individuals (represented by the restrictor of the quantifier) to create a new predicate of individuals (plural or atomic). This new predicate is true of individuals that do not overlap with the individual introduced

[2]Similar exceptive-additive expressions have been reported to exist in many other languages (e.g. German (Mayer 1993), Hebrew (Sevi 2008), Yiddish, Hebrew, Estonian, Ukrainian, Lithuanian, Nanai, Kalmyk, and Ossetian (Oskolskaya 2011), Russian, Turkish, Hindi, Bulgarian, Persian (Vostrikova 2019b), Mandarin (Liu 2019)) We will solely focus on English *besides*, leaving open the possibility of extending the proposed account to other languages.

[3]We only explicitly discuss *but* in what follows and assume that what is said about it carries over to *except*. There are, however, differences. For instance, *except* unlike *but* must be able to receive a clausal analysis (e.g. García Álvarez 2008, Vostrikova 2019a, 2021, Mayr and Vostrikova 2023.)

in the exceptive phrase as proposed by Hirsch (2016).[4] The containment inference is treated as a definedness condition. *But* presupposes that every atomic individual among the individual(s) it introduces is a member of the restrictor set.

(11) $[\![but]\!]^{g,w} = \lambda X_e.\lambda f_{<et>}\colon \forall z[z \preceq X \rightarrow f(z) = T].\lambda Y_e.f(Y) = T \,\&\, X \oslash Y$

With these assumptions, the restrictor set resulting for (10) is as shown in (12).[5]

(12) $[\![student\ but\ Ann]\!]^{g,w} = \lambda y_e.y$ is a student$_w$ & $y \oslash$ Ann
defined only if Ann is a student$_w$

The prejacent of the *EXH* operator obtains the meaning shown in (13). It represents a universal quantificational claim with Ann subtracted from the domain.

(13) $[\![IP]\!]^{g,w} = \text{T}$ iff $\forall x[x$ is a student$_w$ & $x \oslash Ann \rightarrow x$ passed$_w]$

The exception inference is contributed by *EXH*, which asserts the prejacent in (13) and excludes alternatives. The alternatives to be used by *EXH* are structural focus alternatives constructed through F-marking. As seen in (10), the presence of *but* obligatorily triggers F-marking on the DP following it, and the resulting alternatives must be utilized by *EXH*. Given this, the alternatives are formed by making a substitution in the prejacent in the position of the F-marking. Following Katzir (2007) and Fox and Katzir (2011), we adopt the assumption that the alternatives are no more complex than the original sentence. Assuming there are four students Ann, Bill, Carl and Dan (assumed in the discussion throughout), the set of structural alternatives is as shown in (14).

(14) $Alt_{str}(IP) \subseteq \left\{ \begin{array}{l} \text{[[every [student [but Ann]]] passed]} \\ \text{[[every [student [but Bill]]] passed]} \\ \text{[[every [student [but Carl]]] passed]} \\ \text{[[every [student [but Dan]]] passed]} \end{array} \right\}$

In contrast to Hirsch (2016) and much of the literature on implicatures, we use an *EXH* operator that can generate contradictions, as defined in (15). This formulation is therefore closer in spirit to Gajewski (2008) and Crnič (2018)—although the former does not employ *EXH* in the strict sense—and the literature on negative polarity expressions (e.g. Chierchia 2013). The choice of a contradiction generating *EXH* is not made for reasons of exposition. Our reasons for this choice will be discussed in Section 4.3. *EXH* asserts its prejacent while negating all alternatives not entailed by it.

(15) $[\![EXH_{ALT}\ \phi]\!]^{g,w} = \text{T}$ iff $[\![\phi]\!]^{g,w} = \text{T}$ & $\forall q \in ALT[\lambda w'.[\![\phi]\!]^{g,w'} \not\Rightarrow q \rightarrow q(w) = \text{F}]$

In the case of (10), *ALT* is the set of propositions corresponding to the sentences in the set of structural alternatives in (14). Formally, this can be accomplished by impos-

[4] We assume that the domain of individuals forms a join-semilattice (Link 1983) structured by the part-of relation. Individual X is a part of individual Y, $X \preceq Y$, iff there is an individual Z such that the individual sum of X and Z, $X \oplus Z$, is identical to Y. Individuals X and Y do not overlap, $X \oslash Y$, iff there is no Z such $Z \preceq X$ and $Z \preceq Y$. An individual without proper parts is called an atom, whereas an individual with proper parts is referred to as a plurality. Capital letters are used for variables over pluralities or atoms. Non-capital letters are reserved for variables over atoms.

[5] We assume that predicates are inherently pluralized (or obligatorily come with Link's 1983 *-operator). Singular and plural marking on nominal predicates add presuppositions of atomicity and non-atomicity on the argument variable, respectively (Sauerland 2003). In actuality, probably only one of them should be viewed as a lexical presupposition, whereas the other would be an anti-presupposition.

ing the condition that the set of alternatives utilized by *EXH* must be a subset of the propositions denoted by the sentences in the set of structural alternatives.[6]

For (10) all alternatives other than the prejacent can be consistently negated by *EXH*. The resulting interpretation for the entire sentence is given in (16): the first conjunct presents the quantificational claim with domain subtraction, while the other conjuncts result from negating the alternatives. Considering that every student who is not Ann passed, along with some student who is not Bill, Carl, or Dan not passing, it follows that Ann must be the student who did not pass.[7] [8]

(16) $[\![(10)]\!]^{g,w}$ = T iff $\forall x[x \in \{\text{Bill, Carl, Dan}\} \rightarrow x \text{ passed}_w]\&$
$\neg\forall x[x \in \{\text{Ann, Carl, Dan}\} \rightarrow x \text{ passed}_w]\&$
$\neg\forall x[x \in \{\text{Ann, Bill, Dan}\} \rightarrow x \text{ passed}_w]\&$
$\neg\forall x[x \in \{\text{Ann, Bill, Carl}\} \rightarrow x \text{ passed}_w]$

This approach explains why *but* can apply to both quantifiers with universal *every* and negative existential *no*, accounting for the first key property of exceptives discussed in section 1: (17a) gets an LF like the one in (17b), fully parallel to the one with *every*.

(17) a. No student but Ann passed. $\rightsquigarrow$ Ann passed
b. [EXH_{Alt} [[no [student [but Ann_F]]] passed]]

The truth-conditions in (18) obtain: if no student who is not Ann passed, and some student who is not Bill, Carl, or Dan passed, it logically follows that Ann must be the student who passed, yielding the desired positive inference.

(18) $[\![(17b)]\!]^{g,w}$ = T iff $\neg\exists x[x \in \{\text{Bill, Carl, Dan}\} \,\&\, x \text{ passed}_w]\&$
$\exists x[x \in \{\text{Ann, Carl, Dan}\} \,\&\, x \text{ passed}_w]\&$
$\exists x[x \in \{\text{Ann, Bill, Dan}\} \,\&\, x \text{ passed}_w]\&$
$\exists x[x \in \{\text{Ann, Bill, Carl}\} \,\&\, x \text{ passed}_w]$

2.2 The semantic treatment of the distribution puzzle

The leading idea in the existing explanation of the second key property of exceptives, namely, their limited distribution, first proposed by von Fintel (1993, 1994), is that it

[6] More precisely:

(i) a. $[\![EXH_{ALT}\ \phi]\!]^{g,w}$ is defined only if $\text{ALT} \subseteq \{\lambda w'.[\![\phi']\!]^{g,w'}: \phi' \in \text{ALT}_{\text{str}}(\phi)\}$
b. $\text{ALT}_{\text{str}}(\phi) = \{\phi': \phi'$ is derived from ϕ by replacing the F-marked α by β such that β is at most as complex as $\alpha\}$

[7] Note that this result obtains regardless of the differing containment presuppositions of the prejacent and its alternatives. For instance, if Bill is not a student and Ann, Carl and Dan are, the negation of the second alternative in (14) says that not all of Ann, Carl and Dan passed, the negation of the third that not both Ann and Dan passed and the negation of the fourth that not both Ann and Carl passed. Taken together this still entails that Ann did not pass while every other student did.

[8] (i) with adjective *smart* modifying *student* suggests that Ann is a smart student, despite *but* only requiring her to be a student under an LF where *but* forms a constituent with *student* to the exclusion of *smart*. If Ann is a non-smart student, however, she would not be in any of the sets in the alternatives in (16). The prejacent would entail all the alternatives and application of *EXH* would be vacuous, which would be in conflict with the non-vacuity requirement discussed in (58) below. Arguably, a speaker uttering (i) must therefore presuppose that Ann is smart.

(i) Every smart student but Ann passed the test. $\rightsquigarrow$ Ann is a smart student who did not pass

is semantic in nature. He suggests that the meaning of these constructions is necessarily contradictory. Crucially, on the *EXH*-based view the contradictions arise from the interaction of the functional elements *but* and *EXH* alone. Following Gajewski (2002) such contradictions are assumed to lead to ungrammaticality. We illustrate this idea through three representative cases: an upward entailing existential, a non-monotonic modified numeral, and a downward entailing modified numeral.

Given the assumptions outlined in the previous subsection, the LF predicted for the existential claim in (19a) is as shown in (19b).

(19) a. *Some student but Ann passed.
b. [EXH_{Alt} [[some [student [but Ann_F]]] passed]]

The resulting truth-conditions in (20) are contradictory. While none of the individual negated alternatives contradicts the existential claim contributed by the prejacent, they are jointly inconsistent with it. They require that none of the students passed, which contradicts the prejacent. With *but* required to be c-commanded by *EXH*, (19b) is the only available LF and ungrammaticality is predicted for (19a).[9]

(20) $[\![(19b)]\!]^{g,w}$ = T iff $\exists x[x \in \{\text{Bill, Carl, Dan}\} \;\&\; x \text{ passed}_w]\&$
$\neg\exists x[x \in \{\text{Ann, Carl, Dan}\} \;\&\; x \text{ passed}_w]\&$
$\neg\exists x[x \in \{\text{Ann, Bill, Dan}\} \;\&\; x \text{ passed}_w]\&$
$\neg\exists x[x \in \{\text{Ann, Bill, Carl}\} \;\&\; \text{x passed}_w]$

Consider now (21a) with a with non-monotonic *exactly two* and its LF in (21b).[10]

(21) a. *Exactly two students but Ann passed.
b. [EXH_{ALT} [[exactly two [students [but Ann_F]] passed]]

The truth-conditions in (22) are again contradictory. Consider a scenario in which only Carl and Dan passed, thereby rendering the prejacent true. This is compatible with the first negated alternative in the second line only if Ann also passed. However, the next two lines cannot be true in this scenario, as there are now exactly two students who passed in both {Ann, Bill, Dan} and {Ann, Bill, Carl}. Parallel considerations apply

[9] When the context involves only two salient students, such as Ann and Bill, only one alternative to the prejacent is available in the case of (19b), and it can be consistently negated together with the prejacent (Gajewski 2013, Hirsch 2016), as shown in (i).

(i) $[\![(19b)]\!]^{g,w}$ = T iff $\exists$x[x$\in${Bill} & x passed_w] & $\neg\exists$x[x$\in${Ann} & x passed_w]

One way of addressing this issue would be to to assume that the use of (19a) in such a situation is ruled out by a principle preventing the use of an existential quantifier when the conditions for using a definite article are satisfied (Hirsch 2016, cf. Heim's 1991's related but different principle of Maximize Presupposition). Somewhat problematically, *but* with a singular definite is not acceptable itself, as shown by (ii). We return to a suggestion for future research with respect to (ii) in section 6.

(ii) *The student but Ann passed.

[10] Note that it is crucial here that (modified) numerals quantify over atomic individuals. Otherwise, issues with double counting would result: if the domain for the quantifier in (ia) contains pluralities and exactly the two atomic students Bill and Carl passed, (ia) would be wrongly predicted to be true given that the plurality Bill⊕Carl also passed. For illustration only we represent this requirement as an atomicity presupposition on the verbal predicate for now. In section 3 we will be more precise about this.

(i) Exactly three students passed.

to the other two negated alternatives in (22). (21a) thus is correctly expected to be unacceptable in such situations.[11]

(22) $[\![(21b)]\!]^{g,w}$ = T iff | {Bill, Carl, Dan} ∩ {y: y passed$_w$} | = 2 &
| {Ann, Carl, Dan} ∩ {y: y passed$_w$}| ≠ 2 &
| {Ann, Bill, Dan} ∩ {y: y passed$_w$}| ≠ 2 &
| {Ann, Bill, Carl} ∩ {y: y passed$_w$}| ≠ 2

Finally consider (23a) with the downward monotone *at most two* and its LF in (23b).

(23) a. *At most two students but Ann passed.
b. [EXH$_{Alt}$ [[at most two [students [but Ann$_F$]]] passed]]

Once again the resulting truth-conditions in (24) are contradictory and unacceptability follows. While the prejacent (given in the first line) is compatible with the first negated alternative in a scenario where Ann, Carl, and Dan all passed, they together will not be compatible with the rest of the negated alternatives. The same is true mutatis mutandis for the other negated alternatives in (24).

(24) $[\![(23b)]\!]^{g,w}$ = T iff $|\{\text{Bill, Carl, Dan}\} \cap \{y : y \text{ passed}_w\}| \leq 2$ &
$|\{\text{Ann, Carl, Dan}\} \cap \{y : y \text{ passed}_w\}| > 2$ &
$|\{\text{Ann, Bill, Dan}\} \cap \{y : y \text{ passed}_w\}| > 2$ &
$|\{\text{Ann, Bill, Carl}\} \cap \{y : y \text{ passed}_w\}| > 2$

2.3 The challenge posed by *besides*

2.3.1 *Besides* is not just domain subtraction

Given the role of *EXH* in accounting for the distribution of *but* just discussed and the contrast in co-occurrence restrictions between *besides* and *but*, a hypothesis regarding the freer occurrence of *besides* could be the following: while *besides* indeed shares the domain subtraction component with *but*, sentences containing *besides* – at least when it occurs with non-universal quantifiers – could lack *EXH* in their LF.[12] This way *besides* should be able to occur with non-universal quantifiers, but there would not be any additive inference coming with such sentences.

This prediction does not align with the evidence, as demonstrated by the tests in (25). If the additive inferences were not obligatory when *besides* is used, the sequences in (25) should be consistent, contrary to what is observed (Vostrikova 2019b).

(25) a. #Ann didn't pass. But some student besides Ann passed.
b. #Ann didn't pass. And at most two students besides Ann passed.

[11] If the size of the domain of the quantifier equals the number of the numeral, *EXH* would lead to a consistent result and thus acceptability: assume of the three students Ann, Bill and Carl only the latter two passed. Without Ann considered, exactly two students passed making the prejacent true. Without Bill or Carl, however, not exactly two students passed making the negation of the alternatives true as well. A parallel issue arises with downward monotone modified numerals. In both cases it could be addressed by letting *EXH* also have access to the alternative where no individual is subtracted. Then there would be an alternative whose negation would be incompatible with the prejacent and the rest of the negated alternatives, as it would require that not exactly two of Ann, Bill, and Carl passed.

[12] On this view *EXH* might actually be present with universal quantifiers as this would do no harm given the discussion above.

In this respect, *besides* contrasts with *other than*, which seems to contribute containment and domain subtraction, but no negative or positive inference.[13]

(26) a. Ann didn't pass. But some student other than Ann passed.
b. Ann passed. Some student other than Ann also passed.

When the second sentences in (25a) and (25b) are embedded under *think*, as in (27a) and (27b) respectively, the additive inference that Ann passed becomes an entailment of the matrix subject's doxastic state, conflicting with the continuation in (27c). The additive inference should therefore not be optional.

(27) a. Edgar thinks that some student besides Ann passed, ...
b. Edgar thinks that at most two students besides Ann passed, ...
c. #but he is not sure if Ann passed.

The puzzle posed by *besides* is that it interacts with universal quantifiers in a manner similar to *but*, seemingly suggesting a strengthening effect through *EXH*. However, the analysis of *but*-exceptives cannot be straightforwardly applied to *besides*, as it would incorrectly predict that *besides* should also be ungrammatical with non-universal quantifiers like *at most one* and *some*, just as *but* is.

2.3.2 Previous treatments of *besides*

Two kinds of approaches to exceptive-additive *besides* can be found in the literature. The first assumes no ambiguity in the semantics of *besides*, while the second does. None of them fully meets the challenges posed by the EAG in (8).

The first type by von Fintel (1989) provides a unified treatment of *besides* with universal quantifiers and the non-monotonic quantifier *exactly n*. Here *besides* like any exceptive contributes domain subtraction. In addition the minimality condition requires that the subtracted set is the set with the minimal cardinality that has to be subtracted in order for the quantificational claim to be true. The lexical entry proposed is shown in (28): *besides* composes with a set denoted by the DP it introduces[14], a set denoted by the restrictor, the determiner and the set denoted by the scopal argument. It subtracts the first set from the restrictor of the quantifier and adds the minimality condition (underlined in (28)).

(28) $[\![besides]\!]^{g,w} = \lambda C_{<et>}.\lambda A_{<et>}.\lambda D_{<<et><et,t>>}.\lambda P_{<et>}.D(A-C)(P)\&$
$\underline{\forall Y[|C| > |Y| \rightarrow \neg D(A-Y)(P)]}$

The work of the minimality condition can be seen with (29a), which obtains the interpretation shown in (29b). The first conjunct asserts that removing Ann from the domain results in exactly two students passing. The second states that subtracting a set with a smaller cardinality than that of {Ann} will render the *exactly*-claim false. There is only one set with a smaller cardinality than {Ann} – the empty set. Subtracting the empty set has no effect, and we can simplify the truth-conditions as in (29c), requiring that not exactly two students passed overall. If exactly two students, other than Ann, passed, and the total number of students who passed isn't exactly two, then it follows that Ann must be a student who passed. This is because including Ann in the

[13] We thank an anonymous reviewer for suggesting this contrast.

[14] Note that this approach assumes that the DP following an exceptive is a set denoting expression.

domain cannot reduce the number of passing students. Therefore, combining the minimality condition with the initial quantificational claim involving domain subtraction successfully captures the additive inference the sentence in (29a) comes with.

(29) a. Exactly two students besides Ann passed. $\rightsquigarrow$ Ann passed

b. $[\![(29a)]\!]^{g,w}$ = T iff $|(\{z : z \text{ is a student}_w\} - \{\text{Ann}\}) \cap \{x : x \text{ passed}_w\}| = 2$ &
$\underline{\forall Y[|\{\text{Ann}\}| > |Y| \rightarrow |(\{z : z \text{ is a student}_w\} - Y) \cap \{x : x \text{ passed}_w\}| \neq 2\ ]}$

c. $[\![(29a)]\!]^{g,w}$ = T iff $|(\{z : z \text{ is a student}_w\} - \{\text{Ann}\}) \cap \{x : x \text{ passed}_w\}| = 2$ &
$\underline{|\{z : z \text{ is a student}_w\} \cap \{x : x \text{ passed}_w\}| \neq 2\ ]}$

This treatment applies to cases involving universal and negative quantifiers as well, and it successfully explains why the inference is negative in the former and positive in the latter. As noted by von Fintel (1989) it does, however, encounter a challenge when extended to upward entailing numerals, as *at least two* in (30). The truth-conditions predicted are contradictory: they require, on the one hand, that with the singleton containing Ann subtracted from the domain of students at least two students passed and, on the other hand, that with the empty set subtracted fewer than two students passed. Again, the result is contradictory and (30) should therefore be degraded. Notice that this reasoning extends to all upward entailing quantifiers, and von Fintel's approach therefore also makes incorrect predictions for *besides* with existentials. In such cases the prejacent will, abstractly, say that by subtracting the individual following *besides* from the domain there is guaranteed to be a plurality X making the scope argument P true but by not subtracting anything no such X can be found. Again, not both requirements can be satisfied at the same time.

(30) At least two students besides Ann passed.

von Fintel's work also proposes a direct comparison with *but*-exceptives. The difference between the entry for *but* in (31) and the one for *besides* in (28) lies in the underlined condition: while *but*-exceptives share their subtractive semantics with *besides*, the underlined condition (dubbed the leastness condition in the literature, cf. Gajewski 2008) is stricter. It quantifies over all possible subtraction sets that do not have the original subtraction set as a subset. It states that such subtractions do not make the quantificational claim true. While this accounts for the compatibility of *but* with *every* and *no*, and its incompatibility with an existential like *some*, it also derives a contradiction with *exactly*. Given the stricter distribution of *but* discussed above, this is a welcome result.

(31) $[\![but]\!]^{g,w} = \lambda C_{<et>}.\lambda A_{<et>}.\lambda D_{<<et><et,t>>}.\lambda P_{<et>}.D(A - C)(P)\&$
$\underline{\forall Y[C \not\subseteq Y \rightarrow \neg D(A - Y)(P)]}$

Consider (32) and a set of students consisting of Ann, Bill, Carl, and Dan, with Bill and Carl being the students who passed. Then the quantificational claim with Ann removed from the domain is true, satisfying the first condition. The second condition states that removing any set that doesn't include Ann will result in falsity. Removing nothing at all (i.e., subtracting $\emptyset$) should make the claim false. This is only possible if Ann passed as well (since the other two students who passed are still in the domain). But then removing Bill or Carl should also make the claim false, which is impossible, as the remaining sets still should contain two passing students.

(32) *Exactly two students but Ann passed.

The second type of approach is exemplified by Vostrikova (2019b). This work treats the exceptive-additive flip as a genuine case of ambiguity, albeit of a structural type. The exceptive *besides* receives the entry in (33a), which is equivalent to the formulation above, but with leastness treated as a presupposition. The additive *besides* receives a parallel treatment, but with an additivity condition replacing leastness as a presupposition. The additive inference with existentials is derived as follows. (4) receives the truth-conditions shown in (34a), with the additivity presupposition stating that any set containing Ann also contains a student who passed. This is only possible if Ann passed.

The fact that the additive reading is not possible with universal quantifiers is derived as follows: the universal with *besides* under the additive interpretation would receive the meaning shown in (34b): any set containing Ann contains every student who passed. This is only possible if all students (including Ann) passed. The presupposition is therefore stronger than the assertion, which states that the universal claim holds for a smaller domain that does not include Ann. This rules this interpretation out. Thus, universals are only possible with the exceptive *besides*. In addition, given the close similarity in meaning between the exceptive and additive *besides*, it is proposed that this ambiguity can be derived structurally.

This analysis, however, cannot be straightforwardly extended to modified numerals. Consider *exactly two*, as in (29a). The additive reading would require that every set containing Ann contain exactly two students who passed. This is not on the right track, since the singleton set *Ann* does not contain two students.

(33) a. $[\![besides_{EXC}]\!]^{g,w} = \lambda C_{<et>}.\lambda A_{<et>}.\lambda D_{<<et><et,t>>}.\lambda P_{<et>}$:
$\underline{\forall Y[C \cap Y \neq \emptyset \rightarrow \neg D(A \cap Y)(P)]}.\ D(A - C)(P)$

b. $[\![besides_{ADD}]\!]^{g,w} = \lambda C_{<et>}.\lambda A_{<et>}.\lambda D_{<<et><et,t>>}.\lambda P_{<et>}$:
$\underline{\forall Y[C \cap Y \neq \emptyset \rightarrow D(A \cap Y)(P)]}.\ D(A - C)(P)$

(34) a. $[\![some\ student\ besides_{ADD}\ Ann\ passed]\!]^{g,w}$ = T iff $\exists x[x \in (\{\text{x: x is a student}\} - \{Ann\})$ & x passed$_w$]
is defined only if $\forall Y[\{Ann\} \cap Y \neq \emptyset \rightarrow \exists x[x \in (\{\text{x: x is a student}\} \cap \{Y\})$ & x passed$_w$]

b. $[\![every\ student\ besides_{ADD}\ Ann\ passed]\!]^{g,w}$ = T iff $\forall x[x \in (\{\text{x: x is a student}\} - \{Ann\}) \rightarrow x$ passed$_w$]
is defined only if $\forall Y[\{Ann\} \cap Y \neq \emptyset \rightarrow \forall x[x \in (\{\text{x: x is a student}\} \cap \{Y\}) \rightarrow x$ passed$_w$]

2.3.3 The plot

In this paper, we advocate a no-ambiguity approach to address the EAG in (8), building on von Fintel's unified approach to the uses of *besides* with *every*, *no* and *exactly* as our starting point. We demonstrate that this approach can be extended to all cases of modified numerals and existential quantifiers by adopting several assumptions with mostly independent theoretical motivations. We will also adopt von Fintel's (1989) idea that *besides*-exceptives differ from *but* in operating over a smaller set of alternative subtractions.

The first assumption involves the syntactic separation of the domain subtraction carried out by *besides* and the negation of alternatives performed by *EXH*, following the account of *but* in section 2.1. We argue that the only difference from exceptive *but* lies in the kind of alternatives used by *EXH*, allowing for straightforward results

akin to those in von Fintel's *exactly*-cases. The second assumption is that the meaning of modified numerals can be decomposed into a degree quantifier and a *many*-operator within its scope, aligning with prior literature on modified numerals (Hackl 2000, Heim 2000a, Mayr and Meyer 2014, Buccola and Spector 2016). The third assumption is that *many* can be transformed into *exactly* using a maximality operator, as proposed by Buccola and Spector (2016). Given that this step creates a syntactic constituent within the scope of a modified numeral with the meaning equivalent to *exactly*, *EXH* associated with *besides* can apply at the level of this constituent, capturing the positive inference for all modified numerals. The fourth assumption is a generalization of this maximalization operator whereby the account can be straightforwardly extended to existential quantifiers and wh-phrases with *besides*.

2.3.4 A caveat: connected vs. preposed *besides*

In addition to so-called connected uses of *besides*-phrases, i.e., uses where it follows the restrictor, it can also occur preposed. In this paper we largely ignore the latter. The reason for this is that there are crucial empirical differences between the two cases that allow us to adopt the idealizing assumption that they should be treated separately. First, the preposed (35c) contributes an anti-containment inference—that is, the inference that the individual introduced by *besides* is not part of the restrictor set of the quantifier—in contrast to the familiar containment inference of the connected (35a). With containment and anti-containment respectively, the unacceptability of (35b) and (35c) follow (under the assumption that Ann but not Bill is a girl).

(35) a. Some girls besides Ann passed the test. ⇝ Ann is a girl and passed
b. #Some girls besides Bill passed the test. ⇝ Bill is a girl and passed
c. #Besides Ann, some girls passed the test. ⇝ Ann is not a girl but passed
d. Besides Bill, some girls passed the test. ⇝ Bill is not a girl but passed

Second, preposed *besides* with universals does not exhibit this pattern: (36a) receives the familiar exceptive reading, while (36b) receives the additive inference. Seemingly, which of these readings surfaces depends on whether containment is satisfied or not, suggesting that neither a containment nor an anti-containment inference obtains here. The connected exceptive in (36c), on the other hand, does exhibit containment.

(36) a. Besides Ann, every girl passed the test. ⇝ Ann did not pass
b. Besides Bill, every girl passed the test. ⇝ Bill passed
c. #Every girl besides Bill passed the test. ⇝ Bill is a girl and did not pass

Our theory potentially provides insight into why the additive reading with containment is not available for *besides* in (35c). According to our account, *besides* must be interpreted below the scope of *some*. If we assume that the surface position corresponds to the position where *besides* is interpreted, the mechanisms proposed here cannot apply in (35c).

Another case we do not address here is preposed *besides* with focus associates, as in (37), which can only have the additive interpretation (Vostrikova 2019b). Vostrikova (2019b) suggests that the preposed instances without containment discussed above are, in fact, analogous to these cases, as *besides* does not operate on the domain of the quantifier but instead behaves similarly to (37c). This direction appears to be promising. In this case, the restriction observed in (35c) essentially is the restriction seen in (38)

(unacceptable under the assumption that Ann is a girl).

(37) a. Besides Ann, Bill$_F$ danced with Mary. ⇝ Ann danced with Mary
b. Besides Ann, Bill danced with Mary$_F$. ⇝ Bill danced with Ann
c. Besides Ann, Bill$_F$ passed the test. ⇝ Ann passed the test

(38) #Ann came. Some girls came too.

These points relate to an issue pointed out to us by a reviewer. They report that preposed cases of *besides* might be similar to seemingly connected cases of *besides* with a prosodic contour that feels parenthetical, which they say is obligatory for them for some of the examples discussed in this paper. Such prosodic patterns definitely warrant future research. Let us point out though that preposed *besides* and parenthetical prosody on connected *besides* are presumably not one and the same thing. As seen from (37) and (38), preposing, on the one hand, affects the inferences arising in non-trivial ways. Connected cases of *besides*, on the other hand, even with a parenthetical prosody, uniformly follow the EAG in (8) regarding the emergence of exceptive and additive readings and moreover reliably yield a containment inference.[15]

3 The proposal

3.1 *Besides* with universal and negative existential quantifiers

Our treatment of *besides* with universal *every*, negative *no*, and non-monotone *exactly-n* quantifiers builds on a number of ideas by von Fintel (1989), implemented here with the help of *EXH* and alternatives following the works cited above: first, parallel to exceptives the core contribution of *besides* is domain subtraction. Second, both negative and positive inferences arise from the negation of alternative quantificational statements varying in domain subtractions. Finally, *besides* makes a more limited set of alternative subtractions available when compared to exceptives. We implement these ideas by drawing on the general approach to exceptives outlined in section 2.1.

We propose that *besides* works like *but*, which was defined in (11): it performs domain subtraction and introduces a containment presupposition:

(39) $[\![besides]\!] = \lambda X_e.\lambda f_{<et>}\colon \forall z[z \leq X \rightarrow f(z) = T].\lambda Y_e.f(Y) = T \,\&\, X \oslash Y$

The exceptive and the additive inferences are contributed by *EXH*. Similar to what was discussed for *but*, *EXH* obligatorily c-commands *besides*. The LF for (40a) therefore looks as in (40b).

(40) a. Every student besides Ann passed. ⇝ Ann did not pass
b. [EXH$_{ALT}$ [$_{IP}$ every [student [besides Ann]] passed]]

The meaning of the prejacent of *EXH* is the same as for the case with *but*. This is the first alternative in (41) saying that every student minus Ann passed, which is only

[15] The same reviewer provides another puzzling case. Preposed *besides* with upward-entailing modified numerals seemingly behaves more similarly to preposing with *every* than with *some* as can be seen from comparing (i) to (35) and (36).

(i) a. Besides Ann, at least two girls passed the test. ⇝ Ann passed
b. Besides Bill, at least two girls passed the test. ⇝ Bill passed

defined if Ann is a student. The key difference to the case with *but* lies in the non-identical alternatives to the prejacent. For (40b) there is only one, namely the second alternative in (41). This is the proposition that all of the students, i.e., including Ann, passed.

(41) $$Alt = \left\{ \begin{array}{l} \lambda w : \text{Ann is a student}_w. \\ \qquad\qquad \forall x[(x \text{ is a student}_w \,\&\, x \oslash \text{Ann}) \rightarrow x \text{ passed}_w] \\ \lambda w. \forall x[x \text{ is a student}_w \rightarrow x \text{ passed}_w] \end{array} \right\}$$

The prejacent does not entail the second alternative in (41). *EXH* therefore asserts the prejacent and negates this second alternative. The truth-conditions for the sentence are as in (42). If every student who is not Ann passed, but some student did not pass, then Ann is the student who did not pass. This correctly captures the negative inference (40a) comes with. Again, the inference of containment is contributed by *besides* at the presuppositional level.

(42) $$[\![(40b)]\!]^{g,w} = \text{T iff } \forall x[(x \text{ is a student}_w \,\&\, x \oslash \text{Ann}) \rightarrow x \text{ passed}_w] \,\&\, \neg\forall y[y \text{ is a student}_w \rightarrow \text{y passed}_w]$$
$$\text{defined only if Ann is a student}_w$$

These assumptions alone do not deliver the correct truth-conditions for a sentence where the DP following *besides* is plural, as in (43a) or (43b). The issue arises because the negation of the only alternative, essentially *every student passed*, is compatible with only one of the two individuals that constitute the plurality passing. In fact, however, the exception inference applies to both individuals.

(43) a. Every student besides Ann and Bill passed.
$\rightsquigarrow$ Ann and Bill both did not pass

b. Every student besides the twins passed. $\rightsquigarrow$ the twins both did not pass

In order to capture this fact, we make the additional assumption that a plural DP comes with its own set of alternatives. Hence, the list of alternatives will include the alternatives where the plurality is replaced by the DPs denoting the plurality subparts as shown in (44).[16][17]

(44) $$Alt = \left\{ \begin{array}{l} \lambda w : \text{Ann and Bill are students}_w.\ \forall x[(x \text{ is a student}_w \,\&\, x \oslash (\text{Ann} \oplus \text{Bill}) \rightarrow x \text{ passed}_w] \\ \lambda w : \text{Ann is a student}_w\ .\ \forall x[(x \text{ is a student}_w \,\&\, x \oslash \text{Ann}) \rightarrow x \text{ passed}_w] \\ \lambda w : \text{Bill is a student}_w\ .\ \forall x[(x \text{ is a student}_w \,\&\, x \oslash \text{Bill}) \rightarrow x \text{ passed}_w] \\ \lambda w. \forall x[x \text{ is a student}_w \rightarrow x \text{ passed}_w] \end{array} \right\}$$

With (44), *EXH* yields the truth-conditions in (45) for (43a). They capture the negative inference of (43a) correctly. The prejacent together with the negation of the second alternative in (44) entails that Bill did not pass. The same reasoning applied to the third alternative guarantees that Ann did not pass.

(45) $$[\![(43a)]\!]^{g,w} = \text{T iff } \forall x[(x \text{ is a student}_w \,\&\, x \oslash (\text{Ann} \oplus \text{Bill})) \rightarrow x \text{ passed}_w] \,\&\, \exists y[(y \text{ is a student}_w \,\&\, y \oslash \text{Ann}) \,\&\, \neg y \text{ passed}] \,\&$$

[16] Note that semantically alternatives of the form *Every student including Ann passed* or *Every student including Bill passed* are Strawson-equivalent to the one with *including Ann and Bill*.

[17] For (43b) this assumption necessitates access to the atoms of the plurality denoted by *the twins*. If the twin plurality is contextually salient, such alternatives might be available in line with Katzir's 2007 assumptions about substitution for structural alternatives mediated by the context.

$\exists z[(z \text{ is a student}_w \ \& \ z \oslash \text{Bill}) \& \neg z \text{ passed}] \ \&$
$\exists x[x \text{ is a student}_w \ \& \ \neg x \text{ passed}_w]$
defined only if Ann and Bill are students$_w$

To simplify the discussion, we will focus solely on cases where the individual introduced by *besides* is atomic from this point onward.

Consider now (46a) with the LF in (46b) and the alternatives for the prejacent of *EXH* in (46c).

(46) a. No student besides Ann passed. $\rightsquigarrow$ Ann passed
b. [EXH$_{ALT}$ [$_{IP}$ no [student [besides Ann]] passed]]
c. $Alt = \left\{ \begin{array}{l} \lambda w : \text{Ann is a student}_w \ . \ \neg\exists x[(x \text{ is a student}_w \ \& \ x \oslash \text{Ann}) \ \& \ x \text{ passed}_w] \\ \lambda w. \neg\exists x[x \text{ is a student}_w \ \& \ x \text{ passed}_w] \end{array} \right\}$

The truth-conditions are as in (47), which entails that Ann is a student who passed: if no student other than Ann passed, but there is a student who did pass overall, that student must be Ann. This captures the positive inference associated with (46a).

(47) $[\![(46b)]\!]^{g,w}$ = T iff $\neg\exists y[(y \text{ is a student}_w \ \& \ y \oslash \text{Ann}) \ \& \ y \text{ passed}_w] \ \&$
$\exists x[x \text{ is a student}_w \ \& \ x \text{ passed}_w]$
defined only if Ann is a student$_w$

3.2 Options for generating the alternatives

While we will not commit to a particular way of generating the semantic alternatives in (41) and (46c), we want to point out that there are different ways that would in principle be compatible with our goals and the present setting. Recall once more that we adopted Katzir's 2007 structural approach to alternatives.

One option would be to assume that the placement of focus is what differentiates between *but* and *besides*. Instead of F-marking the DP following *besides*, as in the case of *but* in (48a), and forming alternatives by substituting this DP for one of at most the same complexity, we F-mark *besides* itself, as in (48b).

(48) a. [EXH$_{ALT}$ [$_{IP}$ every [student [but Ann$_F$]] passed]]
b. [EXH$_{ALT}$ [$_{IP}$ every [student [besides$_F$ Ann]] passed]]

If the alternative to *besides* were *including*, defined for instance as in (49), we could arguably get the semantic alternatives in (41) for the prejacent of *EXH* in (48b).

(49) $[\![\textit{including}]\!] = \lambda X_e.\lambda f_{<et>}\colon \forall z[z \leq X \rightarrow f(z) = T].\lambda Y_e.f(Y) = T$

Notice that this rests on the idea that *students including Ann* is equivalent to *students* modulo the presupposition that Ann is a student. That is, *including* would be redundant in the assertive component.[18]

[18] An anonymous reviewer points out that *including* comes with its own restrictions. Specifically, Zuber (1998) notes that inclusion phrases are not compatible with left downward monotone quantifiers.

(i) *Few/no students including John passed.

(i) might then suggest that *including* is not redundant to the extent suggested by (49) and it might not be straightforward to use it as an alternative to *besides* for our goals. In response one could, however, explore if the lexical contribution of *including* to the assertive component might not be redundant after all and that the degradedness of (i) is the result of strengthening required because of this very redundancy.

The second option would be to say that the entire exceptive phrase is F-marked, as in (50). If we moreover required that the non-identical alternatives generated through this F-mark have to be structurally simpler, we would end up with *Every student passed* as an alternative. The net result would, again, be the semantic alternatives in (41). This second option might be preferable over the first one because it is not obvious that there is obligatory F-marking on *besides*.

(50) [EXH_{ALT} [$_{IP}$ every [student [besides Ann]$_F$] passed]]

A third option brought to our attention by a reviewer could be to not make a difference in terms of F-marking between *but* and *besides*. That is, also *besides* might force F-marking of the DP following it, as in (51). In contrast to *but*, however, the alternatives would have to be parts of the denotation of the DP. If, crucially, there were a zero-individual (Bylinina and Nouwen 2018), which is part of every individual, again the semantic alternatives in (41) could follow. This hinges not only on there being a zero-individual in the domain but also some of way of referring to it linguistically.

(51) [EXH_{ALT} [$_{IP}$ every [student [besides Ann$_F$]] passed]]

There might well be other options. We leave further exploration of this issue to future research.

3.3 *Besides* and *exactly*

Given what was said, (52a) with *exactly* gets the LF shown in (52b) and the alternatives to the prejacent of *EXH* are as in (52c).

(52) a. Exactly one student besides Ann passed.
b. [EXH_{ALT} [$_{IP}$ [exactly one student [besides Ann]] passed]]
c. $Alt_{\text{str}}(IP) = \left\{ \begin{array}{l} \text{exactly two students besides Ann passed} \\ \text{exactly two students including Ann passed} \end{array} \right\}$

We adopt the semantics for *exactly* shown in (53a). This function first composes with a degree, then with a predicate of plural individuals (its restrictor), then the resulting function composes with the second predicate of individuals (the scopal argument). It returns truth if and only if there exists a plurality that satisfies both predicates, has a cardinality of the degree provided by the numeral, and no plurality with a larger cardinality satisfies the same conditions.[19] This is implemented in (53a) by utilizing the metalinguistic *max* function defined in (53b), which applies to a predicate of degrees and returns the maximal degree satisfying that predicate if a degree satisfying the degree predicate exists and otherwise degree 0 (cf. Hackl (2000) for a related treatment with *max*).[20] In *exactly n*, *max* applies to the predicate of degrees that selects

This strengthening might then be ignored by the exhaustification applying in the case of *besides* whereby, effectively, *including* would not make a contribution to meaning when used in an alternative to *besides*. Alternatively, one might look for other lexical alternatives to *besides* performing the role required for our purposes.

[19] We assume without argument that numerals denote numbers and that numbers are a type of degree (e.g. Hackl 2000, Nouwen 2008 a.o.). Alternatively, numbers could be seen as some form of individuals or entities with their own denotation type.

[20] The latter clause is needed to account for the *exactly zero* case, when there is no relevant plurality, thus the set of degrees is empty (Buccola and Spector 2016).

cardinalities of the pluralities satisfying both the restrictor and the scope of *exactly n*. *Exactly n* states that the highest cardinality satisfying these conditions is *n*.

(53) a. $[\![exactly]\!] = \lambda d_d.\lambda f_{<et>}.\lambda g_{<et>}.max(\lambda d'_d.\exists X[|X| = d' \ \& \ f(X) = g(X) = \mathrm{T}]) = d$

b. $max(P_{<dt>}) = \iota d[P(d) = \mathrm{T} \ \& \ \forall d'[P(d') = \mathrm{T} \rightarrow d' \leq d]]$ if $\exists d[P(d) = \mathrm{T}]$,
0 otherwise

The prejacent of *EXH* in (52b) gets the following truth-conditions: it is true if and only if there are precisely two students, other than Ann, who passed.

(54) $[\![IP]\!]^{g,w}$=T iff max($\lambda d.\exists X[|X|$=d & X is a student$_w$ & $X \oslash$Ann & X passed$_w$])=2

EXH returns truth-conditions in (55): there is precisely one student, excluding Ann, who passed, but not exactly one student passed in total. Since including Ann in the group of students couldn't have reduced the total number of passing students compared to when Ann was excluded, this can only hold if Ann indeed passed.

(55) $[\![(52b)]\!]^{g,w}$= T iff max($\lambda d.\exists X[|X|$=d & X a student$_w$& $X \oslash$Ann & X passed $_w$])=2
& $max(\lambda m.\exists Y[|Y| = m$ & Y is a student$_w$ & Y passed $_w$]) $\neq 2$
defined only if Ann is a student$_w$

This discussion makes the difference between exceptive constructions and *besides* clear. As demonstrated in section 2.2, exceptive *but* is incompatible with *exactly* due to a distinct method of constructing alternatives. Specifically, in exceptive constructions, the element marked with focus is the DP following the exceptive marker. Hence, the alternatives vary with respect to the elements subtracted from the domain of the quantifier. The absence of such alternatives in the case of *besides* is crucial for the proposed account of the interaction between *besides* and *exactly*.

3.4 Upward entailing numerals and *besides*

3.4.1 The challenge of extending the approach

Recall from section 2.3.2 that von Fintel's 1989 original approach runs into a problem with upward entailing numerals such as in (56). It predicts them to be degraded. All else being equal, this problematic prediction is shared by the present implementation in terms of *EXH*. The reason for this is evident: if the only alternative in (56b) different from the prejacent to *EXH* in (56a) is also entailed by it, *EXH* does not negate it.

(56) At least two students [besides Ann] passed. $\rightsquigarrow$ Ann passed

(57) a. [EXH$_{ALT}$ [$_{IP}$ [at least two students [besides Ann]] passed]]

b. $Alt_{\mathrm{str}}(IP) = \left\{ \begin{array}{l} \text{at least two students besides Ann passed} \\ \text{at least two students including Ann passed} \end{array} \right\}$

In order to rule out the LF in (57a) regardless, we appeal to a constraint restricting the occurrence of *EXH* when it does not contribute to the meaning of the sentence (Spector 2013, Hirsch 2016, Gajewski 2013, Fox and Spector 2018).

(58) NON-VACUITY: *EXH[A]* is infelicitous if *EXH[A]* is equivalent to *A*.

But ruling out LF (57a) alone is insufficient given that the sentence in (56) is perfectly grammatical. Thus some LF for it must be licensed. Moreover *besides* contributes an

additive meaning here. According to the present perspective this must involve an *EXH* somewhere in the LF. That is, we only want to rule out vacuous *EXH*.

3.4.2 The decompositional approach to modified numerals

Our solution to this puzzle is based on the idea that modified numerals are quantifiers over degrees, requiring quantifier raising for their interpretation (see Hackl 2000, Mayr and Meyer 2014, Buccola and Spector 2016). This theory extends the idea of treating degree constructions quantificationally, as originally suggested in Heim 2000a, to modified numerals. When such a quantifier moves, it can optionally leave behind a constituent with the meaning *exactly*. Because of this we can directly apply our approach for *exactly n* numerals to these cases by placing *EXH* in the scope of the upward entailing numeral itself.

Following Hackl (2000), we assume that *at least* when supplied a degree d is a quantifier over degrees with the semantics given in (59): like any degree quantifier, it seeks to compose with a predicate of degrees, an expression of type $< dt >$. It introduces existential quantification over degrees and returns truth if there is a degree that satisfies this predicate that is equal to d or greater.[21]

(59) $[\![at\ least]\!] = \lambda d_d.\lambda f_{<dt>}.\exists d'[d' \geq d\ \&\ f(d') = \mathrm{T}]$

From this it follows that *at least two* in (60a) does not compose with the plural predicate *students* directly. We adhere to the conventional degree semantics assumption that a degree quantifier is a sister of a silent operator *many* of type $< d, << et >, < et >, t >>$. Due to the type mismatch, *at least two* in (60a) cannot be interpreted in its initial position. It undergoes quantifier raising, leaving behind a trace of type d as in the LF in (60b). A numerical index, matching the index on the trace, is then merged below the landing site of *at least two*.

(60) a. At least two students passed.
b. [at least two [$_{IP}$ 1 [many d_1 students passed]]]

The definition of *many* is given in (61). It is a function that first composes with a degree and subsequently with two predicates of individuals and returns truth if and only if there is a plurality that satisfies both predicates with the cardinality of the degree. The trace left by *at least two* serves as the first argument of the silent *many*.

(61) $[\![many]\!] = \lambda d_d.\lambda f_{<et>}.\lambda g_{<et>}.\exists X[|X| = d\ \&\ f(X) = g(X) = \mathrm{T}]$

The sister of *at least two* in (60b) denotes the predicate of degrees in (62a). Accordingly, the overall truth-conditions are as in (62b), as desired.

(62) a. $[\![IP]\!]^{g,w} = \lambda d_d.\exists X[|X| = d\ \&\ X \text{ is a student}_w\ \&\ X \text{ passed}_w]$
b. $[\![(60b)]\!]^{g,w} = \mathrm{T}$ iff $\exists d[d \geq 2\ \&\ \exists X[|X| = d\ \&\ X \text{ is a student}_w\ \&\ X \text{ passed}_w]]$

3.4.3 Turning *many* into *exactly*

The decompositional analysis creates an additional position for the *EXH* associated with *besides* in the scope of the modified numeral above *many*, as in (63). This alone is

[21] Hackl (2000) extends this approach to *exactly n*-quantifiers. In principle this would also be compatible with the present proposal, but we do not discuss this here.

not sufficient to account for the interaction of *at least two* and *besides*. Because of the fact that *many* is upward-entailing, *EXH* would still be vacuous in (63).

(63) [at least two [1 [EXH_{ALT} [many d_1 students besides Ann passed]]]]

Following Buccola and Spector (2016), we suggest that a silent operator transforms *many* into *exactly*. This operator is inserted in order to prevent a violation of the non-vacuity constraint (58) which would prohibit *EXH*. Because the resulting constituent has an exact-interpretation, it will be possible to derive the additive inference with modified numerals in full analogy to cases with *exactly*. According to Buccola and Spector (2016) modified numerals can undergo quantifier raising twice, as illustrated in (64). The first movement is already familiar from the previous discussion. The trace d_2 left behind by the second movement is type-shifted through the operator *MAX* in (65). It employs the metalinguistic *max* function used above for *exactly* and defined in (53b). *EXH* is positioned above the first landing site of the modified numeral.

(64) [at least one [2 [$_{IP_2}$ EXH_{ALT} [$_{IP}$ [MAX d_2] [1 [many d_1 student besides Ann passed]]]]]]

(65) $[\![MAX]\!] = \lambda d_d.\lambda P_{<dt>}.max(P) = d$

With these assumptions the denotation of IP is equivalent to that of an *exactly*-statement: excluding Ann, the number of students who passed is exactly the degree $g(2)$.[22]

(66) $[\![IP]\!]^{g,w}$ = T iff $max(\lambda d.\exists X[|X| = d$ & X is a student_w & $X \oslash \text{Ann}$ & X $\text{passed}_w])$ $= g(2)$

EXH can be non-vacuously applied to (66). To spell it out, the sole alternative distinct from the prejacent IP is given in (67).

(67) [[MAX d_2] [1 [many d_1 student including Ann passed]]]

EXH asserts the prejacent and negates this alternative yielding (69). This says that without Ann, the number of students who passed is exactly the degree $g(2)$ but the overall number of students who passed is not $g(2)$. This gives rise to the additive inference that Ann passed.

(68) $[\![IP_2]\!]^{g,w} = max(\lambda d_d.\exists X[|X| = d$ & X is a student_w & $X \oslash \text{Ann}$ & X $\text{passed}_w])=g(2)$ & $max(\lambda d.\exists X[|X| = d$ & X is a student_w & X $\text{passed}_w]) \neq g(2)$

With abstraction over degree $g(2)$ in (68) and *at least two* applied to the resulting predicate of degrees, the truth-conditions of the sentence in (56) on the LF in (64) is as in (69). They yield truth if and only if there exists a degree equal to 2 or larger such that this is the number of students who passed when Ann is not taken into consideration, yet it is not the total number of students who passed. The inference that Ann passed is imported from (68). Consider a situation where Ann did not pass. Then the number of students who passed remains unchanged, regardless of whether we include Ann. Thus, the truth-conditions will not be satisfied in this scenario.

[22] This outcome could possibly also be achieved with an additional *EXH* instead of *MAX*, similar to the upper bound meaning for bare numerals such as *two students*. The alternatives for this *EXH* would be constructed by substituting the trace. This might necessitate the existence of a null individual. Otherwise, numerals with an existential semantics would come with an existence entailment, which would run into problems with downward entailing numerals. See also footnote 25.

(69) $[\![(64)]\!]^{g,w}$ = T iff $\exists d[d \geq 2 \& max(\lambda d'.\exists X[|X| = d' \ \& \ X \text{ is a student}_w \ \& \ X \oslash \text{Ann} \ \& \ X \text{ passed}_w]) = d \ \& \ max(\lambda d'.\exists X[|X| = d' \ \& \ X \text{ is a student}_w \ \& \ X \text{ passed}_w]) \neq d]$

Notice that there is always an exact number of people who passed. When we choose to describe this number with the modified numeral *at least two*, we are not providing this number directly—presumably for reasons of ignorance—but describing it indirectly via comparison with another number, namely 2.

3.5 *Besides* and downward entailing numerals

This decompositional account is readily extended to downward entailing modified numerals like *fewer than n* and *at most n*. We treat them as existential quantifiers over degrees, as illustrated in (70), following Buccola and Spector (2016).

(70) a. $[\![\textit{fewer than}]\!] = \lambda d_d.\lambda f_{<dt>}.\exists d'[d' < d \ \& \ f(d') = \text{T}]$
b. $[\![\textit{at most}]\!] = \lambda d_d.\lambda f_{<dt>}.\exists d'[d' \leq d \ \& \ f(d') = \text{T}]$

As Buccola and Spector (2016) note, *MAX* plays a crucial role even in the absence of *besides* as in (71), in contrast to upward entailing modified numerals. The assumption that $max(\varnothing) = 0$ avoids an undesirable existential entailment for (71) that some students passed. The truth-conditions for (71) are as in (72). They yield true when two students or fewer passed and false when three or more did. The delimitative aspect typically associated with the modified numeral itself is integrated into the maximality function. When three or more students passed, the maximal cardinality for which a plurality of students who passed exists is not smaller or equal to 2.[23]

(71) At most two students passed.

(72) $[\![(71)]\!]^{g,w}$ = T iff $\exists d[d \leq 2 \ \& \ max(\lambda d'.\exists X[|X| = d' \ \& \ X \text{ is a student}_w \ \& \ X \text{ passed}_w]) = d]$

The example in (73a) with *besides* would have the LF in (73b).

(73) a. At most two students besides Ann passed.
b. [at most two [2 [$_{IP}$ EXH$_{ALT}$ [[MAX d_2] [1 [many d_1 student besides Ann passed]]]]]]

EXH yields a denotation for *IP* saying that the number of students who passed without considering Ann equals degree $g(2)$, and $g(2)$ is not the number of students who passed overall. Again, this carries the additive inference that Ann passed.

(74) $[\![IP]\!]^{g,w} = max(\lambda d.\exists X[|X| = d \ \& \ X \text{ is a student}_w \ \& \ X \oslash \text{Ann} \ \& \ X \text{ passed}_w])$=g(2) $\& \ max(\lambda d.\exists Y[|Y| = d \ \& \ Y \text{ is a student}_w \ \& \ Y \text{ passed}_w]) \neq g(2)$

The overall truth-conditions for (73a) are given in (75). They say that there exists a number, which represents the maximal count of students who passed without consid-

[23] With *MAX* present the existential semantics for the downward entailing modifiers in (70) yield truth-conditions equivalent to those obtained through Hackl's 2000 treatment as negative quantifiers, as in (i). That is, with (70) the presence of *MAX* is crucial. Buccola and Spector's 2016 reasons for adopting (70) is connected to the behavior downward entailing modifiers show with collective predicates. Cf. section 5.

(i) a. $[\![\textit{fewer than}]\!] = \lambda d_d.\lambda f_{<dt>}.\neg\exists d'[d' \geq d \ \& \ f(d') = \text{T}]$
b. $[\![\textit{at most}]\!] = \lambda d_d.\lambda f_{<dt>}.\neg\exists d'[d' > d \ \& \ f(d') = \text{T}]$

ering Ann, and this number is distinct from the total number of students who passed, and it is less than or equal to 2. The additive inference of Ann passing is inherited from (74). Consider for instance a scenario where Ann did not pass, and the total count of passing students is 1. Given (75), (73a) is false in this context. This is because the number of students who passed is unaffected by considering Ann or not. Consequently, there is no degree that satisfies (74) with abstraction over $g(2)$.

(75) $[\![(73a)]\!]^{w,g}$ = T iff $\exists d[d \leq 2$ & $max(\lambda d'.\exists X[|X| = d'$ & X is a student$_w$ & $X \oslash$ Ann & X passed$_w$]) = d & $max(\lambda d'.\exists Y[|Y| = d'$ & Y is a student$_w$ & Y passed$_w$]) $\neq d]$

Downward entailing numerals of the form *fewer than n* are treated in a parallel manner. To illustrate, the predicted truth-conditions for (76a) with the familiar LF in (76b) are as in (77). As the reader can verify, these truth-conditions necessitate that Ann passed, accurately capturing the additive inference.

(76) a. Fewer than two students besides Ann passed. $\rightsquigarrow$ Ann passed
b. [fewer than two [2 [$_{IP}$ EXH$_{ALT}$ [[MAX d$_2$] [1 [many d$_1$ student besides Ann passed]]]]]]

(77) $[\![(76b)]\!]^{w,g}$ = T iff $\exists d[d < 2$ & $max(\lambda d'.\exists X[|X| = d'$ & X is a student$_w$ & $X \oslash$ Ann & X passed$_w$]) = d & $max(\lambda d'.\exists Y[|Y| = d'$ & Y is a student$_w$ & Y passed$_w$]) $\neq d]$

In contrast to upward entailing modified numerals combining with *besides*, placing *EXH* above a downward entailing modified numeral as shown in (78) (where we are glossing over the internal composition in the scope of the modified numeral) is non-vacuous. This is to be expected given our discussion for *besides* with *no* above.

(78) [EXH$_{ALT}$ [fewer than two students student [besides Ann] passed]]

The predicted truth-conditions are given in (79): if we do not count Ann, fewer than two students passed; if we count her, two or more students passed. In other words, (76a) on the LF in (78) is predicted to be equivalent to the sentence in (80).

(79) $[\![(78)]\!]^{w,g}$ = T iff
$\exists d[d < 2$ & $max(\lambda d'.\exists X[|X| = d'$ & X is a student$_w$ & $X \oslash$Ann & X passed$_w$]) = $d]$
& $\neg\exists d[d < 2$ & $max(\lambda d'.\exists X[|X| = d'$ & X is a student$_w$) & X passed$_w$]) = $d]$

(80) Exactly one student besides Ann passed.

The sentence in (76a) does not require that exactly two students passed. It is true if only one student passed, namely Ann. The truth-conditions in (77) allow for this, unlike (79). This means that the LF in (76b) must be available. Whether (78) is also available is difficult to tell given that its truth-conditions entail those of (76b). It would be interesting to investigate if the LF in (78) could be ruled out in favor of the weaker (76b) through some form of competition. Luka Crnič (pc.) suggests (81) as a test case. Here the second disjunct would entail the first one when interpreted using the weak LF (76b) but not when interpreted with the strong LF (78). Hurford's generalization says that a disjunction is blocked if one disjunct entails the other (cf. Chierchia et al. 2012 a.o.). So on this reasoning, if (81) is acceptable, it should involve LF (76b).

(81) Fewer than two students besides Ann passed or only Ann did.

A possible complication is that truth-conditions similar to those of (78) could also be

obtained with local *EXH* for *besides* as in (76b) and a further higher *EXH* yielding an indirect scalar implicature for *fewer than two* to the effect that at least one student who is not Ann passed. We must leave this issue for future research.

3.6 *Besides* with existential quantifiers

As discussed and shown again in (82), *besides* can also occur with existential *some* and *several*. Crucially, additive inferences are licensed here as well.[24]

(82) a. Some student(s) besides Ann passed. ⇝ Ann passed
b. Several students besides Ann passed. ⇝ Ann passed

Placing *EXH* above the existential will not produce the desired result for the reason discussed in section 2.3.2. The account for *besides* with modified numerals discussed above can, however, be generalized to the cases in (82) with some tweaks. For this we adopt a *MAX* that can apply to both individuals and degrees and predicates thereof, as in the entry in (83).

(83) $[\![MAX]\!] = \lambda x_{\sigma}.\lambda P_{<\sigma t>}.max(P) = x$, where $\sigma = e$ or $\sigma = d$

Moreover, the metalinguistic function *max* is generalized accordingly as in (84).[25]

(84) $$max(P) = \begin{cases} \iota d[P(d) = \mathrm{T} \ \&\ \forall d'[P(d') = \mathrm{T} \ \rightarrow d' \leq d]], & \text{if } P \in D_{<dt>} \ \&\ \exists d[P(d) = \mathrm{T}], \\ 0, & \text{if } P \in D_{<dt>} \ \&\ \neg\exists d[P(d) = \mathrm{T}] \\ \iota X[P(X) = \mathrm{T} \ \&\ \forall Y[P(Y) = \mathrm{T} \ \rightarrow Y \leq X]], & \text{if } P \in D_{<et>} \ \&\ \exists X[P(X) = \mathrm{T}], \\ \text{undefined} & \text{otherwise} \end{cases}$$

The LF for (82a) with singular then looks as in (85). Here the restrictor of the existential quantifier is interpreted in its base-position, composed by predicate modification with the verbal predicate. *MAX* takes the resulting predicate of individuals and the trace of the quantificational determiner as its arguments.[26] The quantificational determiner is optionally interpreted as in (86).

(85) [some 2 [EXH_{ALT} $[_{IP}$ [MAX t_2] [student besides Ann passed]]]]

(86) $[\![some]\!] = \lambda f_{<et>}.\exists X[f(X) = T]$

The denotation of *IP* says that $g(2)$ is identical to the maximal atomic student not overlapping with Ann who passed. This is equivalent to saying that $g(2)$ is the unique student who is not Ann and passed.

[24] We think that what we say below can be extended to bare plurals with *besides* as well, which generally also give rise to additive inferences. Given the ongoing debate about how to best analyze bare plurals, however, we leave investigation of this to future research.

[25] The two cases in (84) do not work in a fully parallel fashion. If there is no entity making P true, it yields 0 in the case of degrees and an undefined value in the case of individuals. The two cases could be brought into line if a zero individual were assumed, i.e., a full lattice with a bottom element were assumed for the domain of individuals (cf. Landman 2011, Buccola and Spector 2016, Bylinina and Nouwen 2018). The two cases could then even be collapsed if degrees and individuals formed the domain of entities.

[26] These are simplified assumptions about the LF. In parallel to the case of modified numerals one could assume that there is an equivalent to a silent *many* in the base position. Since this would amount to another existential quantifier over individuals, we leave it out to avoid clutter. Also, it is conceivable that (90) is the result of the copy-theory of movement, in which case there would also be a copy of the restrictor next to the quantifier.

(87) $[\![IP]\!]^{w,g} = max(\lambda y.y \in [\![\textit{student besides Ann}]\!]^w \ \&\ y \text{ passed}_w) = g(2)$

EXH now negates the alternative with *including Ann*. In order for both the prejacent and the negation of this alternative to stand a chance of being true together there must be unique students who passed in both cases. All else being equal, the prejacent in (87a) amounts to (88a) and its negated alternative to (88b). But for there to be a unique student who passed in both cases exactly one student can have passed overall. This in turn would lead to the requirement that the unique student be identical and different from $g(2)$ at the same time.

(88) a. $max(\{\text{Bill, Carl, Dan}\} \cap \{Y : Y \text{ passed}_w\}) = g(2)$
b. $max(\{\text{Ann, Bill, Carl, Dan}\} \cap \{Y : Y \text{ passed}_w\}) \neq g(2)$

To avoid this problematic prediction we adopt the view that the singular feature on *student* need not be interpreted in the alternative negated by *EXH*.[27] With that the negated alternative looks as in (89). Imagine $g(2)$ is Bill, and Ann and Bill passed. Then the prejacent in (88a) is true, and (89) is true because the maximal plurality that passed here consists of Ann and Bill and $g(2)$ is not identical to that plurality. Notice also that if Ann did not pass, the conflicting requirements on $g(2)$ would resurface. *Some* in (86) simply adds the inference that there is an individual satisfying the discussed requirements. In other words, the additive inference for (82a) with singular is derived.[28]

(89) $max(\{\text{Ann, Bill, Carl, Dan, Ann}\oplus\text{Bill}, \dots\} \cap \{Y : Y \text{ passed}_w\}) \neq g(2)$

The plural case in (82a) and (82b) works in parallel with the LF being (90). The prejacent of *EXH* in (91a) says that $g(2)$ is the maximal student plurality that passed. Assume Bill and Carl passed and $g(2)$ is the corresponding plurality. For the negated alternative in (91b) to be true at the same time Ann must have also passed. Again, the additive inference is derived.

(90) $[_{IP_3}$ some/several 2 [EXH_{ALT} [[MAX t_2] [students besides Ann passed]]]]

(91) a. $max(\{\text{Bill, Carl, Dan, Bill}\oplus\text{Carl}, \dots\} \cap \{Y : Y \text{ passed}_w\}) = g(2)$
b. $max(\{\text{Ann, Bill, Carl, Dan}, \dots, \text{Bill}\oplus\text{Carl}, \dots, \text{Ann}\oplus\text{Bill}\oplus\text{Carl}, \dots\}) \neq g(2)$

3.7 *Besides* in wh-questions

Besides also occurs with wh-phrases. In that case, too, it contributes additive inferences.

(92) a. Who besides Ann passed? $\rightsquigarrow$ Ann passed

[27] The idea that certain features can be ignored in the interpretation of alternatives has been motivated among other things by indexical and plurality features on bound pronouns in parallelism environments (e.g. Heim 2008, Sauerland 2013 building on an idea by B. Spector).

[28] On this view (82a) on the singular reading entails that exactly one student who is not Ann passed. Indeed, this exactly-one inference is quite prominent. A reviewer, however, challenges this on the basis of cases like (i). To the extent that (i) is acceptable, though, it need not speak in favor of an account where the exactly-one inference is the result of a scalar implicature. If the domain of students is kept to Ann and one other student for the first sentence in (i) and enlarged to include all students for the second sentence, felicity would be expected on the present account for (i). Notice that upon hearing the second sentence in (i) there is an intuition that the first one gets reparsed.

(i) Some student besides Ann passed. In fact, all twenty of them did.

b. Which student(s) besides Ann passed? $\rightsquigarrow$ Ann passed

The wh-phrase is standardly assumed to scope over the question operator *?*. Since *?* returns a question denotation (Hamblin 1973, Karttunen 1977), the wh-phrase taking scope over the result as in (93) is composed with such a denotation. If *EXH* were to apply globally, it would have to be generalized accordingly.[29]

(93) [[wh [-o/students besides Ann]] 2 [? [t_2 passed]]]

Consider, however, LFs for (92a) and for (92b) with plural in (94a) and (94b), respectively. These LFs are modeled on those for existential quantifiers with *besides* in section 3.6. Here and below we assume *-o* to denote the set of singularities and pluralities of human individuals.[30]

(94) a. [wh- 2 [? [$_{IP}$ EXH_{ALT} [[MAX t_2] [[-o besides Ann] passed]]]]]
b. [wh 2 [? [$_{IP}$ EXH_{ALT} [[MAX t_2] [[students besides Ann] passed]]]]]

At the *IP*-level (94a) and (94b) are fully parallel to the LF in (90) for plural *some* with *besides*. In the case of (94b), for instance, it says that $g(2)$ is identical to the maximal student plurality not overlapping with Ann that passed and that $g(2)$ is not the maximal student plurality overall that passed. This requires that both $g(2)$ and Ann passed. *?* is standardly defined as in (95a) on Karttunen's 1977 account. It takes a proposition and returns a set thereof, i.e., it creates a question denotation. (95b) gives Heim's 2000b semantics for wh-question adopted to the present setting with the restrictor in the question nucleus. Crucially, it takes a question denotation where abstraction over an individual has occurred and existentially closes the predicate.[31]

(95) a. $[\![?]\!] = \lambda p_{st}.\lambda q_{st}.q = p$
b. $[\![wh]\!] = \lambda f_{\langle e,\langle st,t\rangle\rangle}.\lambda p_{st}.\exists x[f(x)(p) = T]$

The result for (94b) then is as in (96). Each of the propositions in this set says about some student plurality not overlapping with Ann that it is the maximal one that passed but not the maximal one overall. That is, each proposition requires Ann to have passed.

(96) $[\![(94b)]\!]^g = \lambda p.\exists X[p = \lambda w'.max([\![\textit{students besides Ann}]\!]^{w'} \,\&\, \text{passed}_{w'}) = X \,\&\, max([\![\textit{students including Ann}]\!]^{w'} \,\&\, \text{passed}_{w'}) \neq X]$

$$= \left\{ \begin{array}{l} \lambda w'.max([\![\textit{students besides Ann}]\!]^{w'} \,\&\, \text{passed}_{w'}) = \text{Bill}\oplus\text{Carl} \,\&\, \\ \quad max([\![\textit{students including Ann}]\!]^{w'} \,\&\, \text{passed}_{w'}) \neq \text{Bill}\oplus\text{Carl} \\ \lambda w'.max([\![\textit{students besides Ann}]\!]^{w'} \,\&\, \text{passed}_{w'}) = \text{Bill}\oplus\text{Dan} \,\&\, \\ \quad max([\![\textit{students including Ann}]\!]^{w'} \,\&\, \text{passed}_{w'}) \neq \text{Bill}\oplus\text{Dan} \\ \ldots \end{array} \right\}$$

Given the discussion in section 3.6, a fully parallel result obtains for (92b) with singular

[29] In a previous version of this paper, we suggested that *EXH* applies to the weak exhaustive answer to a wh-question. Given the generalization of *MAX* in (83), however, this is not necessary anymore it seems.

[30] This is justified by the fact that an answer to *Who passed?* naming a singularity where in fact more than one person passed is felt to be incomplete (e.g. Groenendijk and Stokhof 1984, Dayal 1996).

[31] LFs like the ones in (94) and the general treatment of wh-questions with the wh-restrictor in the scope of *?* are motivated by Groenendijk and Stokhof's 1984 so-called *de dicto* interpretations of wh-questions. The present account is incomplete insofar as the propositions in the set in (96) do not presuppose that the individual abstracted over is a student. This could be brought about by having the definite description *the students identical to X* with *X* abstracted over instead of simply *X* in the propositions (cf. Rullmann and Beck 1998, Fox 2002, Heim 2019).

on *student*, except that here a uniqueness rather than a mere maximality inference arises. Assuming now that utterance of a wh-question, all else being equal, comes with the speaker presupposition that at least one of its elements is true (e.g. Dayal 1996), it follows that utterance of (92a) and (92b) effectively presupposes that Ann passed.

A reviewer draws our attention to *how many*-questions with *besides* such as (97a), which similarly have additive inferences. Cresti (1995) argues that *how many*-questions allow for an interpretation where the restrictor is equally interpreted below *?* (also see Longobardi 1987, Fox 1999). The LF in (97b) yields the question-equivalent to what we assumed for modified numerals above. Ignoring *EXH*, (98b) asks for the number n such that n is the maximal number of student pluralities who do not include Ann and all of them passed. *EXH* can be inserted just below *?* contributing that this number is different when Ann is included, giving rise to the additive inference.[32]

(97) a. How many students besides Ann passed? $\rightsquigarrow$ Ann passed
b. [wh 2 [? [EXH_{ALT} [[MAX 2_d] [1 [d_1-many students besides Ann passed]]]]]]

4 Discussion of theoretical choices and comparison

4.1 Comparison with *but*

We now briefly check that the assumptions about local application of *EXH* for modified numerals and existentials do not negatively affect the predictions made for them in section 2.2 when occurring with *but*. Let us start by considering the case of a downward-entailing numeral and *but* in (98a) under the LF in (98b).

(98) a. *At most two students but Ann passed.
b. [at most two [2 [$_{IP}$ EXH_{ALT} [[MAX d_2] [1 [many d_1 student but Ann_F passed]]]]]]

The denotation of IP is as in (99). As discussed in Section 2.2, it is contradictory, at least for any number $g(2)$ greater than 0. If $g(2)$ is the number of students who passed excluding Ann, and this is not the total number of students who passed overall (i.e., Ann passed as well), it is impossible for all three alternative sets—those in which Bill, Carl, and Dan are subtracted—to each yield a number of students who passed that is different from $g(2)$. (99) is only consistent if the value assigned to $g(2)$ is 0 and Ann is the sole student who passed.

(99) $[\![IP]\!]^{g,w}$ = T iff
$max(\lambda d.\exists X[|X| = d \;\&\; X \text{ is a student}_w \;\&\; X \oslash \text{Ann} \;\&\; X \text{ passed}_w]) = g(2)$ &
$max(\lambda d.\exists X[|X| = d \;\&\; X \text{ is a student}_w \;\&\; X \oslash \text{Bill} \;\&\; X \text{ passed}_w]) \neq g(2)$ &
$max(\lambda d.\exists X[|X| = d \;\&\; X \text{ is a student}_w \;\&\; X \oslash \text{Carl} \;\&\; X \text{ passed}_w]) \neq g(2)$ &
$max(\lambda d.\exists X[|X| = d \;\&\; X \text{ is a student}_w \;\&\; X \oslash \text{Dan} \;\&\; X \text{ passed}_w]) \neq g(2)$

On the present account, 0 is indeed a permissible value for $g(2)$. However, this predicts the sentence to be truth-conditionally equivalent to any sentence of the form *At most n students but Ann passed*. Following Buccola and Spector (2016:p. 165), we assume that a pragmatic constraint blocks such LFs. With *EXH* embedded under *at least two* in

[32] Interpretation below *?* in such cases is motivated through relative scope of the *d-many* phrase and a modal expression.

(100) what was just said about (99) carries over. Since (98) is contradictory for degrees greater than zero and *at least two* introduces existential quantification over degrees greater than 1, (100) comes out ungrammatical.

(100) *At least two students but Ann passed.

Consider now singular *some* with *but* and the LF in (101b) with embedded *MAX*.

(101) a. *Some student but Ann passed.
b. [some 2 [EXH_{ALT} [[MAX t_2] [student but Ann_F passed]]]]

In order for the singularity presupposition in the prejacent of *EXH* in (102a) to be satisfied only one Bill, Carl and Dan must have passed. Moreover, in order for the negated alternatives in (102b) to (102d) to have a chance at being true at least one must have passed. Given the requirements on the prejacent this student must be Ann. Assume that $g(2)$ is Bill. Then (102a) is true, and so are all of (102b) to (102d). The reason is that $g(2)$ is neither identical to Ann nor the plurality of Ann and Bill. (101a) should come out as acceptable.

(102) a. $max(\{\text{Bill, Carl, Dan}\} \cap \{Y : Y \text{ passed}_w\}) = g(2)$
b. $max(\{\text{Ann, Carl, Dan, Ann}\oplus\text{Carl, } \dots\} \cap \{Y : Y \text{ passed}_w\}) \neq g(2)$
c. $max(\{\text{Ann, Bill, Dan, Ann}\oplus\text{Bill, } \dots\} \cap \{Y : Y \text{ passed}_w\}) \neq g(2)$
d. $max(\{\text{Ann, Bill, Carl, Ann}\oplus\text{Bill, } \dots\} \cap \{Y : Y \text{ passed}_w\}) \neq g(2)$

In order for (101a) to be ruled out nonetheless alternatives where pluralities are subtracted must be allowed as well. For instance, the negation of the alternative where the plurality of Ann and Dan is subtracted in (103) cannot be true together with the prejacent in (102a) and the rest of the negated alternatives in the considered scenario: if $g(2)$ is Bill and Bill is the unique student who is not Ann and passed, the same must hold in (103). With such alternatives the unacceptability of (101a) is predicted even on the LF in (101b) then.[33]

(103) $max(\{\text{Bill, Carl, Bill}\oplus\text{Carl}\} \cap \{Y : Y \text{ passed}_w\}) \neq g(2)$

The only relevant difference between (104) with plural and (101a) with singular is then that in the former $g(2)$ in the prejacent must be a plurality. Otherwise everything is the same and (104a) is ruled unacceptable in the same way.

(104) *Some students but Ann passed.

Given the discussion of *besides* in wh-questions in section 3.7 and the LFs assumed there it also immediately follows that wh-questions with *but* such as in (105) should not be acceptable as information-seeking questions.[34] Remember that the level at which

[33] This, of course, necessitates that a plurality like Ann⊕Dan is an alternative to the atom Ann, which raises non-trivial questions for a theory of alternatives in terms of structural complexity. In particular, some form of referring to the former that is not more complex than *Bill* would have to be available.

[34] It is worth noting that *but*-exceptives are acceptable in rhetorical questions (see (i)) (Horn and Bayer 1984). Here, we follow von Fintel (1993) in assuming that this is possible because rhetorical questions convey the same meaning as negative quantifiers. Thus, the question in (i) has the same meaning as *No one but a total idiot would do this*.

(i) Who but a total idiot would do this?

EXH applies in the case of wh-questions is fully parallel to the one in (104). The result for existential quantifiers with *but* therefore automatically carries over.

(105) a. *Who but Ann passed?
b. *Which student(s) but Ann passed?

4.2 Separation of domain subtraction and the negation of alternatives

The syntactic separation of domain subtraction and the negation of alternatives (Gajewski 2008, 2013, Hirsch 2016, Crnič 2021) is crucial for our account of *besides* with upward entailing quantifiers. This is so because in such examples, *MAX* must structurally intervene between *besides* (and *many*) on the one hand and *EXH* on the other hand. *MAX* creates a non-upward entailing constituent, and only then *EXH* can be non-vacuously applied.

A number arguments has been put forward in support of separating domain subtraction from the negation of alternatives in the analysis of exceptives like *but*. Two of them can be extended to *besides*. First, strong NPIs like *in years* are licensed in the scope of negative quantifiers with *but* as in (106). If the *but*-phrase were to contribute, in addition to domain subtraction, the negation of alternatives, the scope of the negative quantifier would not constitute a downward-entailing (DE) environment (Gajewski 2008). This is because negating a negative existential claim results in a positive existential claim, placing the NPI in an upward-entailing environment.[35]

(106) No one but John has visited me in years.

Empirically, this can be appreciated by looking at the entailment pattern for the whole sentence and observing that the position of *in years* is not in a DE environment globally, as (108a) does not entail (108b).

(107) a. No one but John has visited me last year.
b. No one but John has visited me last spring.

Following Gajewski (2008), strong NPI licensing in the scope of *no one* in (107a) can be accounted for by separating domain subtraction and exclusion of alternatives syntactically, as shown in (108). Although the NPI is not in a downward entailing environment globally, there is a constituent, namely the sister of *EXH* that is downward entailing. This is sufficient for NPI licensing.

(108) [EXH_{Alt} [no one but John has visited me in years]]

In exactly the same way, strong NPIs are licensed in the scope of *no one* whose domain is modified by *besides*, as shown in (109).

(109) No one besides John has visited me in years.

Just as with *but* in (106), the NPI is not in a downward entailing environment overall given that (110a) does not entail (110b). Separating domain subtraction from the negation of alternatives through an LF fully parallel to (108), however, yields a downward entailing constituent.

[35] Strong NPI are sensitive to the presuppositional content of the licensing environment (Gajewski 2011). So even if negation of the alternatives happened at the presuppositional level, this should interfere in the licensing of the NPI in (106).

(110) a. No one besides John visited me last year.
b. No one besides John visited me last spring.

Consider for the second argument, Crnič's 2021 observation of a contrast between (111b) and (111c) when following (111a): while a sentence containing an exceptive cannot be conjoined with another sentence that directly contradicts the exceptive inference, it is acceptable to conjoin it with a modalized version of that sentence.

(111) a. Every student except Dina arrived, ...
b. #and even she did!
c. and perhaps even she did!

Crnič (2021) proposes that this fact can be derived under the assumption that a silent modal operator contributing speaker certainty can intervene between *EXH* and the universal statement. Accordingly, the resulting reading is: 'The speaker is certain that every student who is not Dina arrived, and they consider it possible that not every student who is not Ann arrived, and they consider it possible that not every student who is not Bill arrived ...'. This captures the compatibility of (111a) and (111c).

(112) [EXH_{Alt} [ASSERT [every student except Dina arrived]]]

Similarly, modal suspension readings are available for *besides*: while (113a) is incompatible with a statement that directly contradicts the exceptive inference in (113b), it is compatible with the same statement when embedded under *perhaps*. If an LF similar to (112) is available this is predicted.

(113) a. Every student besides Dina arrived, ...
b. #and even she did!
c. and perhaps even she did!

In summary, our decision to separate the two steps – domain subtraction and negation of alternatives – in the syntax for the exceptive-additive *besides* is well-founded.[36]

4.3 The choice of the *EXH*-operator

We employed a version of *EXH* that can generate contradictions. This type of *EXH* is prevalent in theories concerned with the licensing of polarity items (Chierchia 2013, Spector 2013 a.o.). In contrast, Hirsch (2016) adopts a contradiction-free version of *EXH* for exceptive *but*. This version of *EXH* works with Fox's 2007 notion of innocent inclusion and is used for scalar implicatures and free-choice phenomena.

The contradiction-free EXH^{IE} operator based on innocent exclusion is defined in

[36] NPI *any* is generally assumed to be an existential quantifier. All else being equal, *any* with *but* in (ia) should therefore yield a constituent with a trivial meaning below negation. (ia) should be just as degraded as (ib). By having *EXH* scope above negation in (ia), this can be avoided (Gajewski 2008).

(i) a. I did not eat any vegetables but cucumbers.
b. *I ate some vegetables but cucumbers.

This argument is not applicable to *besides*. The reason is that it can occur with both *any* and *some* as in (ii).

(ii) a. I did not eat any vegetables besides cucumbers.
b. I ate some vegetables besides cucumbers.

(114). Crucially, it does not negate all non-entailed alternatives, but only those that belong to every maximal set of alternatives that can be negated consistently with the assertion of the prejacent.

(114) a. $[\![EXH^{IE}_{ALT}\ \phi]\!]^{g,w} = [\![\phi]\!]^{g,w} = \text{T}\ \&\forall q[q \in \text{IE}(\lambda w'.[\![\phi]\!]^{g,w'}, ALT) \rightarrow q(w) = F]$

b. IE(p, ALT) = $\cap$\{ALT' $\subseteq$ ALT: ALT' is a maximal subset of ALT s.t. ALT'$^{\neg}$ $\cup$ \{p\} is consistent\}

c. ALT'$^{\neg}$ = $\{\neg p' : p' \in ALT'\}$

Now, adopting the contradiction-free EXH^{IE} in (114) to deal with the distribution of *but* can only make a difference relative to an account in terms of our non-contradiction-free *EXH* in cases where the latter derives a contradiction to predict unacceptability. Existential quantifiers with *but* as in (115) are a case in point. EXH^{IE} does not yield a contradiction for (115). Rather its application is vacuous here. Hirsch (2016) therefore suggests that the unacceptability of (115) is an instance of a violation of the non-vacuity constraint (58).[37]

(115) *Some student but Ann passed.

A reviewer draws our attention to cases like (116a).[38] The problem arises in all situations where the size of the subtracted set equals the size of the remaining domain. The intensionalized meaning of the prejacent of EXH^{IE} with our usual four students available is given in (116b).

(116) a. *Some student but Ann and Bill passed.

b. λw.$\exists$x[x$\in$\{Carl, Dan\} & x passed$_w$]

The list of alternatives includes sentences in which one or two individuals are subtracted, each of which is of equal or lesser complexity than the original. Their meanings are shown in (117). The alternative of interest is (117a). The negation of this alternative is compatible with the prejacent regardless of which other alternatives are negated, because the domains of the two existential quantifiers are disjoint.[39]

(117) a. $\lambda w.\exists x[x \in \{\text{Ann, Bill}\}\ \&\ x\ \text{passed}_w]$

b. $\lambda w.\exists x[x \in \{\text{Bill, Dan}\}\ \&\ x\ \text{passed}_w]$

c. $\lambda w.\exists x[x \in \{\text{Bill, Carl}\}\ \&\ x\ \text{passed}_w]$

d. $\lambda w.\exists x[x \in \{\text{Ann, Dan}\}\ \&\ x\ \text{passed}_w]$

e. $\lambda w.\exists x[x \in \{\text{Ann, Carl}\}\ \&\ x\ \text{passed}_w]$

f. $\lambda w.\exists x[x \in \{\text{Ann, Carl, Dan}\}\ \&\ x\ \text{passed}_w]$

[37] Things could be different if Bar-Lev and Fox's 2020 modification of contradiction-free EXH^{IE} were adopted. In addition to innocent exclusion, their version also makes use of innocent inclusion. That is, their operator EXH^{IE}_{II} seeks to include as many alternatives that cannot be negated due to innocent exclusion as possible without making arbitrary choices. Such inclusion could therefore make the application of EXH^{IE}_{II}, if one were to adopt it in the context of *but*, non-vacuous. Indeed, (115) would be such a case. The interpretation predicted for (115) would be that at least two students passed. This prediction is obviously wrong.

[38] In an earlier version of this paper we had adopted EXH^{IE}.

[39] Notice that no issue arises with (i). Here *EXH* and EXH^{IE} both consistently negate all alternatives—essentially the corresponding universal alternatives to the existential ones in (117)—yielding truth-conditions saying that Ann and Bill didn't pass and all other students did.

(i) Every student but Ann and Bill passed. $\rightsquigarrow$ Ann and Bill didn't pass

g. $\lambda w.\exists x[x \in \{\text{Ann, Bill, Dan}\} \,\&\, x \text{ passed}_w]$
h. $\lambda w.\exists x[x \in \{\text{Ann, Bill, Carl}\} \,\&\, x \text{ passed}_w]$

To convince ourselves of this, let us examine the possible maximal sets of alternatives whose negation is consistent with the assertion of the prejacent. Negating any alternative entails that no one in the domain of that alternative passed. Since the prejacent asserts that either Carl or Dan passed, there are two maximal sets of alternatives that can be negated consistently with the prejacent: one containing alternatives that exclude Carl from their domain, shown in (118a), and one containing alternatives that exclude Dan from their domain, shown in (118b). The alternative in which only Ann and Bill remain in the domain appears in both sets. Because there is an innocently excludable alternative, application of EXH^{IE} would not be vacuous. Consequently, the sentence is predicted to be well-formed saying that some student who is not Ann and Bill passed, and Ann and Bill did not pass.

(118) a. $\left\{ \begin{array}{l} \lambda w.\exists x[x \in \{\text{Ann, Bill}\} \,\&\, x \text{ passed}_w] \\ \lambda w.\exists x[x \in \{\text{Bill, Dan}\} \,\&\, x \text{ passed}_w] \\ \lambda w.\exists x[x \in \{\text{Ann, Dan}\} \,\&\, x \text{ passed}_w] \\ \lambda w.\exists x[x \in \{\text{Ann, Bill, Dan}\} \,\&\, x \text{ passed}_w] \end{array} \right\}$

b. $\left\{ \begin{array}{l} \lambda w.\exists x[x \in \{\text{Ann, Bill}\} \,\&\, x \text{ passed}_w] \\ \lambda w.\exists x[x \in \{\text{Bill, Carl}\} \,\&\, x \text{ passed}_w] \\ \lambda w.\exists x[x \in \{\text{Ann, Carl}\} \,\&\, x \text{ passed}_w] \\ \lambda w.\exists x[x \in \{\text{Ann, Bill, Carl}\} \,\&\, x \text{ passed}_w] \end{array} \right\}$

We do not see an easy solution to this problem and conclude that non-contradiction-free *EXH* provides a better fit for the treatment of exceptives and exceptive-additives.[40]

5 Extensions and challenges

5.1 *Besides* and collective predicates

So far, this paper has examined examples involving *besides* with distributive verbal predicates. We observe, however, that examples with modified numerals and *besides* are not compatible with collective predicates like *lift the piano together*:

(119) a. *Ten students besides John lifted the piano together yesterday.
b. *More than ten students besides John lifted the piano yesterday.
c. *Exactly ten students besides John lifted the piano yesterday.
d. *Fewer than ten students besides John lifted the piano yesterday.

Our approach as it stands does not capture the data in (119). Consider (119d). If *MAX* is inserted into the structure as shown in (120a), we obtain a well-formed meaning. The resulting truth-conditions are provided in (120b). These yield true in two types of scenarios. In the first type, the largest group of students who lifted the piano without John consisted of fewer than 10 people. Additionally, there was a distinct event where a larger group of students, including John, lifted the piano together. In the second type

[40] The issue arising with (116a) is a generalization of the problem mentioned in footnote 9 regarding (115) in contexts with only two students. The problem created by (116a) is more severe and cannot be straightforwardly resolved by the competition mechanism discussed there. Possible options for future investigation include competition between *some* and *both* or *all n*, where *n* is a numeral.

of scenario, only a group with John lifted the piano. In this case, the degree that makes the existential quantification true is 0.

(120) a. [[fewer than ten] [2 [EXH_{ALT} [[MAX d_2] [1 [d_1 many students besides John lifted the piano yesterday]]]]]]
b. $[\![(120a)]\!]^{g,w}$ = T iff $\exists d[d < 10$ & $max(\lambda d'.\exists X[|X| = d'$ & X is a student $_w$ & $x \oslash$ John & X lifted$_w$ the piano$]) = d$ &
$max(\lambda d'.\exists Y[|Y| = d'$ & Y is a student $_w$ & Y lifted$_w$ the piano$]) \neq d]$

5.1.1 Modifying maximality

It is known that downward-entailing numerals behave differently depending on whether the predicate in the scope of the numeral is distributive or collective (Winter 2001, Ben-Avi and Winter 2003, Buccola 2015, Spector 2014, Buccola and Spector 2016). (121a), with the distributive predicate *came*, exhibits the familiar upper-bound reading: it means that there is no plurality of students who came with a cardinality greater than five. (121a) can even be true if no students came at all.

(121) a. Fewer than ten students came.
b. Fewer than ten students lifted the piano together yesterday.

(121b) with a collective predicate differs from (121a) in two respects: First, no upper-bound limit is imposed—the sentence is compatible with the possibility that a larger group of students (containing, say, twelve students) also lifted the piano; second, the sentence carries the existential inference that some students did lift the piano.

To account for this an alternative theory of maximality based on logical strength has been proposed (Buccola and Spector 2016, von Fintel et al. 2014, Dayal 1996, Rullmann 1995, Fox and Hackl 2006, Beck and Rullmann 1999, Schlenker 2012). Buccola and Spector (2016) in particular suggest to redefine the maximality function *max* from (53b) as in (122), stated here so as to be applicable to both degrees and individuals in light with our discussion in section 3.6. With regard to a degree d, (122) defines when d counts as maximal relative to a property f of degrees in a given world w. Crucially, the first case in (122) allows for two degrees to both be maximal relative to f in w, as long as neither makes f true in strictly fewer worlds than the other. Again, a special amendment is required to account for the fact that distributive predicates with downward entailing numerals are compatible with situations where no individual satisfying the predicate exists. The second case in (122) covers this.

(122) $max_{inf}(f_{<s<\sigma t>>})(w_s)(g_\sigma)$ = T if $\begin{cases} f(w)(g) = \text{T} \;\&\; \neg\exists h[f(w)(h) = \text{T} \;\&\; \\ \quad \{w' : f(w')(h) = \text{T}\} \subsetneq \{w' : f(w')(g) = \text{T}\}] \text{ or} \\ g = 0 \;\&\; \neg\exists d'[f(w)(d') = \text{T}] \end{cases}$
where $\sigma = e$ or $\sigma = d$

Consequently, redefining *MAX* from (83) in terms of max_{inf}, as done in (123) and inserting it in the scope of the downward entailing numeral as in (124a) and (124b) for (121a) and (121b), respectively, ensures that the resulting interpretation is upper-bounded in the former case only.

(123) $[\![MAX_{inf}]\!]^w = \lambda x_\sigma.\lambda P_{<s<\sigma t>>}.max_{inf}(P)(w)(x)$ = T, where $\sigma = e$ or $\sigma = d$

(124) a. [fewer than ten $[_{IP_1}$ 2 [[MAX_{inf} d_2] [1 d_1 many students came]]]]

b. [fewer than ten [$_{IP_2}$ 2 [[MAX$_{inf}$ d$_2$] [1 d$_1$ lifted the piano together]]]]

The set of degrees for which the denotation of IP_1 in (124a) returns truth in a world w will always be a singleton; its sole member is the degree corresponding to the maximal number of comers in w. In contrast, the predicate denoted by IP_2 need not be a singleton. If, in w, two groups of students consisting of 9 and 10 people lifted the piano, (125b) characterizes the set containing degrees, 9 and 10. This is because both of them are maximal under the definition of maximality given in (122).

(125) a. $[\![IP_1]\!]^{g,w} = \lambda d.max_{inf}(\lambda d'.\lambda w'.\exists X[|X| = d' \ \&\ X \text{ is a student}_{w'} \ \&\ X \text{ came}_{w'}])(w)(d) = \text{T}$

b. $[\![IP_2]\!]^{g,w} = \lambda d.max_{inf}(\lambda d'.\lambda w'.\exists X[|X| = d' \ \&\ X \text{ is a student}_{w'} \ \&\ X \text{ lifted}_{w'} \text{ the piano together}])(w)(d) = \text{T}$

The truth-conditions for (121a) in (126a) thus yield falsity if two groups of students came: one consisting of 9 people and one consisting of 10. This is because the total and thus maximal number of comers is 19, i.e., greater than 10. In contrast to this, the truth-conditions for (121b) in (126b) yield truth in a parallel scenario where two groups consisting of 9 and 10 people lifted the piano. This is because with max_{inf} both 9 and 10 are maximal numbers of piano-lifters.[41]

(126) a. $[\![(121a)]\!]^{g,w} = \text{T}$ iff $\exists d[d < 10 \ \&\ max_{inf}(d)(\lambda d'.\lambda w'.\exists X[|X| = d' \ \&\ X \text{ is a student}_{w'} \ \&\ X \text{ came}_{w'}])(w)(d) = \text{T}]$

b. $[\![(121b)]\!]^{g,w} = \text{T}$ iff $\exists d[d < 10 \ \&\ max_{inf}(\lambda d'.\lambda w'.\exists X[|X| = d' \ \&\ X \text{ is a student}_{w'} \ \&\ X \text{ lifted}_{w'} \text{ the piano together}])(w)(d) = \text{T}]$

5.1.2 Modified maximality and its consequences for *besides*

We now explore the predictions our approach to *besides* makes when combined with this theory. First, consider (127a) with distributive *came* and its LF in (127b). The predicate denoted by *IP* characterizes a non-empty set of degrees precisely when adding John to the domain increases the number of comers. In a scenario where nine students, including John, came overall there is a number smaller than 10—specifically 8—such that, in the domain that does not include John, 8 is the maximal number of comers. However, in the domain that includes John, 8 is not the maximal number of comers. This means the predictions for downward entailing modified numerals with distributive predicates from section 3.5 are preserved under the modified notion of maximality.

(127) a. Fewer than ten students besides John came.

b. [fewer than 10 [$_{IP}$ 2 [EXH$_{ALT}$ [[MAX$_{inf}$ d$_2$] [1 d$_1$ many students besides John came]]]]]

(128) $[\![IP]\!]^{g,w} =$
$\lambda d.max_{inf}(\lambda d'.\lambda w'.\exists X[|X|{=}d' \ \&\ X \oslash \text{John} \ \&\ X \text{ is a student}_{w'} \ \&\ X \text{ came}_{w'}])(w)(d) = \text{T} \ \&\ max_{inf}(\lambda d'.\lambda w'.\exists X[|X|{=}d' \ \&\ X \text{ is a student}_{w'} \ \&\ X \text{ came}_{w'}])(w)(d) = \text{F}$

Consider now the crucial case in (129a) with a collective predicate and its LF in (129b).

[41] This leaves open why there is an existential entailment when downward-entailing numerals combine with collective predicates. Since this is a broader problem in the semantics of modified numerals and is not directly relevant to the interaction between numerals and *besides*, we leave it for future research and refer the reader to Buccola and Spector 2016.

For the time being, let's set aside a possible situation where no group of students that does not include John lifted the piano. Then we cannot find a value for $g(2)$ (other than 0) that can make the prejacent of *EXH* with the meaning (130a) and the negation of its only alternative in (130b) simultaneously true in a given world.

(129) a. Fewer than ten students besides John lifted the piano together.
b. [fewer than 10 [2 [EXH_{ALT} [[MAX_{inf} d_2] [1 d_1 many students besides John lifted the piano together]]]]]

(130) a. $\lambda w.max_{inf}(\lambda d.\lambda w'.\exists X[|X|{=}d$ & $X \oslash$John & X is a student$_{w'}$ & X lifted$_{w'}$ the piano$])(w)(g(2))$ = T
b. $\lambda w.max_{inf}(\lambda d.\lambda w'.\exists X[|X|{=}d$ & X is a student$_{w'}$&X lifted$_{w'}$ the piano$])(w)(g(2))$ = F

Let's consider a possible value of 8 for $g(2)$ and a scenario where the prejacent is true—where a group of students, not including John, consisting of 8 people, lifted the piano together. Given the upward-entailing properties of *many*, if there exists a group of 8 individuals from the set of students without John that lifted the piano together, it logically follows that there is a group of 8 students that lifted the piano together.[42] Now, since the only alternative to the prejacent is entailed by it, *EXH* cannot negate it and applies vacuously violating the non-vacuity condition (58).

Consider now a situation where no group of students without John lifted the piano. In this case, the prejacent and the negated alternative can both be true if there is also at least one group of students including John who lifted the piano. Here, 0 becomes a possible value for $g(2)$ in both clauses in (130). Since the maximal number of students without John who lifted the piano is 0, (130a) is true. Because 0 is not the maximal number of lifters overall, (130b) is also true. All else being equal, (129a) would thus be acceptable in such a scenario. As said above, however, sentences with downward entailing numerals and collective predicates are judged false in situations where no plurality satisfying the predicate exists. Therefore, this scenario is, arguably, ruled out for independent reasons. (129a) thus comes out as unacceptable.

This reasoning extends to the other cases of (modified) numerals with *besides* and collective *lift the piano together* in (119). As just demonstrated with (130), the crucial problem arises at the level of MAX_{inf}. Thus all of (119) are predicted to be degraded.

5.1.3 Further predictions

Some collective predicates allow for downward inferences similar to distributive predicates. For instance, on at least one of its interpretations *gather* allows for an inference from (132a) to (132b), a so-called sub-entailment, similar to what we find in (131) with *come*. This is in contrast to *lift the piano together*, which does not license such a sub-entailment from (133a) to (133b).

(131) a. Ten students came.
b. Nine students came.

(132) a. Ten students gathered downstairs.
b. Nine students gathered downstairs.

(133) a. Ten students lifted the piano together yesterday.

[42] This holds regardless of whether there is a larger group of students of 9 people including John who lifted the piano together. Both 8 and 9 would be considered maximal numbers of students given max_{inf}.

b. Nine students lifted the piano together yesterday.

An entailment between (133a) and (133b), on the one hand, does not hold because max_{inf} allows for more than one maximal degree in this case, as discussed for (125b) above. An entailment between (131a) and (131b), on the other hand, does obtain because here only one maximal degree can exist, as discussed for (125a). On the interpretation with sub-entailments for (132a) and (132b) max_{inf} would similarly allow for only one maximal degree, i.e., an upper-bounded reading can be generated.

Given the possibility of sub-entailments for (132), we expect that on this particular interpretation *besides* is licensed with modified numerals and *gather*. The relative acceptability of the sentences in (134) suggests that this prediction is borne out. An interpretation without sub-entailments more similar to what we find with *lift the piano together* arguably also exists. Here *besides* would not be licensed.

(134) a. ?Ten students besides John gathered downstairs.
b. ?More than ten students besides John gathered downstairs.
c. ?Exactly ten students besides John gathered downstairs.
d. ?Fewer than ten students besides John gathered downstairs.

Moreover, since for us MAX_{inf} is generalized so as to be applicable in the case of existential quantifiers as well, we expect that collective predicates without sub-entailments should also lead to unacceptability in constructions with such quantifiers and *besides*. (135a) and (135b) suggest that this is borne out.

(135) a. *Some students besides John lifted the piano together.
b. *Several students besides John lifted the piano together.

The problem is parallel to what we saw for modified numerals above. With the LF for (135a) and (135b) in (136), *EXH* would have to assert the prejacent in (137a) and the negation of its alternative in (137b). Assume $g(2)$ is a maximal student plurality not including John who lifted the piano together. Then $g(2)$ is also a maximal student plurality who lifted the piano together if John is taken into account. In other words, (136a) and (136b) cannot be true at the same time because the prejacent entails its alternative. So *EXH* is vacuous in (136) and the sentences in (135) come out as unacceptable.

(136) [some/several 2 [EXH_{ALT} [[MAX_{inf} t_2] [students besides John lifted the piano together]]]]

(137) a. $\lambda w.max_{inf}(\lambda X.\lambda w'.X \oslash \text{John} \,\&\, X \text{ is a student}_{w'} \,\&\, X \text{ lifted}_{w'} \text{ the piano}])(w)(g(2)) = \text{T}$
b. $\lambda w.max_{inf}(\lambda X.\lambda w'.X \text{ is a student}_{w'} \& X \text{ lifted}_{w'} \text{ the piano}])(w)(g(2)) = \text{F}$

5.2 *Besides* and upward scale predicates

Our analysis of *besides* occurring with numeral quantifiers crucially relies on the presence of MAX_{inf}. Accordingly, we expect that if this operator cannot be inserted, numerals would not be compatible with *besides*. Now, it has been argued that MAX_{inf} cannot be inserted on top of upward-scale predicates such as *sufficient* in (138a) and certain constructions featuring existential modals, as in (138b). This is because there is no maximal degree that satisfies such predicates (Beck and Rullmann 1999, Fox and Hackl 2006, Buccola and Spector 2016). For example, if five people are sufficient to lift

the piano, then any larger number of people would also be. Likewise, if five people can lift the piano, then it can also be lifted by a larger group of people (Buccola 2016).

(138) a. Five people are sufficient to lift this piano.
b. Five people can lift this piano.

Even though (138a) and (138b) seemingly express the same meaning, we unexpectedly observe a difference in how the two constructions interact with *besides*: while *besides* occurs felicitously with numeral constructions involving *sufficient*, it is not acceptable in constructions that involve possibility modals, as the contrast in (139) shows. The sentence in (139a) means something like this: 'If we put John and any five other people together, the piano can be lifted.' It is appropriate in a context where John has exceptional physical strength and is singled out for this reason. Our conjecture is that the difference lies in the syntax of the two constructions.

(139) a. Five people besides John are sufficient to lift this piano.
b. *Five people besides John can lift this piano.

5.2.1 *Besides* and *sufficient*

Adapting the idea proposed by Beck and Rullmann (1999) for similar sentences, we propose that (138a) can be paraphrased as shown in (140), notably with a predicate *available*. We will take seriously the idea that this predicate is covertly present in the structure of the sentence, a possibility briefly discussed in (Buccola and Spector 2016:p. 175), as in the LF in (141).

(140) 'It is possible (given the physical restrictions of our world) that there are only five people available and the piano gets lifted.'

(141) [[[five people ~~available~~] are sufficient] [PRO_2 to lift this piano]]

Instead of presenting the meaning of *sufficient* as containing existential quantification over possible worlds and a meaning component akin to *only*, as was proposed by Beck and Rullmann (1999), we propose that it introduces universal and existential quantification over possible worlds and composes with two propositional arguments as shown in (142) (here R_{sym} stands for a similarity accessibility relation, R_{phys} represents identity in terms of physical properties).

(142) $[\![\textit{sufficient}]\!]^w = \lambda p.\lambda q.\forall w'[R_{sym}(w')(w) = p(w') = \text{T} \rightarrow \exists w''[R_{phys}(w'')(w') = q(w'') = \text{T}]]$

The first propositional argument of *sufficient* in (141) is the proposition denoted by *five people available*, and the second one is *to lift this piano*. The resulting truth-conditions in (143) capture the idea behind the paraphrase in (140) for (138a): in every world where five people are available (even in the ones where only five people are available), there is a possibility that the piano gets lifted.[43]

(143) $[\![(138a)]\!]^{g,w} = \text{T}$ iff $\forall w'[R_{sym}(w')(w) = \text{T}$ & five people are available in $w' \rightarrow \exists w''[R_{phys}(w'')(w') = \text{T}$ & $g(2)$ lifts the piano in $w'']]$

Combined with our overall assumptions in this paper, the LF for (139a) is as in (144).

[43] We simplify and treat arbitrary *PRO* in (141) and in what follows as a regular free pronoun.

(144) [[five [2 [EXH_{ALT} [[MAX_{inf} d_2] [1 [d_1 many people besides John ~~available~~]]]]] are sufficient] [PRO_2 to lift this piano]]

The truth-conditions are as in (145): in every world where there is a number of 5 or greater, such that this is the maximal number of available people excluding John, but it is not the maximal number of available people, the piano possibly gets lifted.

(145) $[\![(144)]\!]^{g,w}$= T iff $\forall w'$ [$\text{R}_{sym}(w')(w)$ = T & $\exists d[d \geq 5$ &
$max_{inf}(\lambda w'''.\lambda d'.\exists X[|X|=\text{d}' \ \& \ X \oslash \text{John} \ \& \ X \text{ is a person}_{w'''} \& X \text{ is available}_{w'''})$
$(w')(d)$ = T &
$\neg max_{inf}(\lambda w''''.\lambda d'.\exists X[|X|=\text{d}' \ \& X \text{ is a person}_{w''''} \ \& X \text{ is available}_{w''''})(w')(d)$
$\rightarrow \exists w''[\text{R}_{phys}(w'')(w') = \text{T} \ \& \ g(2)$ lifts the piano in $w'']]$

This captures the truth-conditions of the sentence, including the additive inference. Essentially, with these syntactic assumptions, when *besides* appears in a sentence with *sufficient*, it actually occurs inside a clause with a distributive predicate in the scope of the numeral. As a result, this case can be reduced to those we previously analyzed.

5.2.2 *Besides* and *can*

We conjecture that the sentence in (138b), in contrast, does not have a bi-clausal structure. Instead, it has a structure as shown in (146), where the subject is reconstructed below the existential modal. The resulting truth-conditions are then as in (147).

(146) [can [five [2 [[MAX_{inf} d_2] [1 d_1 many people lifted the piano together]]]]]

(147) $[\![(146)]\!]^{g,w}$= T iff $\exists w'[R_{phys}(w')(w)$ = T $\& max_{inf}(\lambda d.\lambda w''.\exists X[|X| = d \ \& \ X$ is a $\text{person}_{w''}$ & X lifted the $\text{piano}_{w''}$ $])(w')(5)$ = T]

Given the LF in (146), to explain the ungrammaticality of *besides* in (139b), we only need to observe that we have already established in section 5.1.2 that there is no place in the prejacent of the modal where *EXH* associated with *besides* can be inserted non-vacuously. This explains why *besides* cannot occur felicitously in this construction.

6 Conclusion and outlook

In this paper, we proposed a unified compositional approach to exceptive and additive uses of *besides*. To our knowledge, this is the first semantic account of *besides* that accurately captures its interaction with universal quantifiers like *every* and *no* as well as all cases of modified numerals, existential quantifiers and wh-expressions. On the present proposal, whether an exceptive or an additive inference arises depends solely on the logical properties of the quantifier involved. More specifically, the exceptive and additive inferences arise from comparing the quantificational statement with domain subtraction performed by *besides* and the alternative statement without this subtraction. The comparison is contributed by *EXH*. For many of the examples it was crucial to apply *EXH* at an embedded level with maximization through MAX_{inf} allowing for non-vacuous exhaustification. Importantly, this account derives the EAG, the empirical generalization mentioned in (8) at the outset of this paper: in all of the environments where exceptive *but* yields ungrammaticality, *besides* is acceptable and leads to an additive inference. In those environments where *but* is acceptable, *besides* is as well and moreover behaves like such a pure exceptive.

Despite this, there remains rich empirical ground for future research. For instance, we have solely focused on the connected uses of *besides*. While there are crucial empirical differences between these and preposed uses of *besides*, as pointed out in section 2.3.4, there are also many similarities that should be accounted for eventually.

Another interesting issue that we have to leave for future research is the interaction between *besides* and quantifiers that are near the endpoints of the scale, such as *many* and *most*. *Besides* can felicitously occur with such quantifiers, as shown in (148). Crucially, (148a) and (148b) are both compatible with Ann having passed and having not passed.

(148) a. Most girls besides Ann passed the test.
b. Many girls besides Ann passed the test.

The data in (148) could be captured by saying that no exceptive or additive inference is derived at all, but they would be equally compatible with an ambiguity analysis. That is, on one interpretation an exceptive interpretation would arise and on another an additive one. Potential support for this latter view comes from the observation that in (149), where only the cardinal interpretation of *many* is available, only the additive reading survives.

(149) There were many girls besides Ann at the gathering.
$\rightsquigarrow$ Ann was at the gathering

The present approach indeed predicts an additive inference for cardinal cases like (149). The predicted truth-conditions can be paraphrased as follows: 'There is a contextually large number n such that, if we exclude Ann, exactly n-many people were at the gathering; but if we include her, the count is no longer exactly n.'

Consider now the other side of the possible source for the ambiguity. For *most* and *many* under the proportional reading, we predict a special kind of exceptive interpretation. Specifically, for (148a), we predict a reading along the following lines: if we remove Ann from the restrictor, it is true that the majority of girls passed the test; however, if we consider the whole domain, it is false that the majority of girls passed. This reading captures the exceptive inference that Ann did not pass, but it also requires that the inclusion or exclusion of Ann makes a crucial difference to the validity of the *most*-claim. Alas, our informants tell us that the sentence does not require this to be the case. Thus, our current proposal does not provide a solution for how the exceptive reading arises for *most* and *many* or how the additive reading with *most* is derived.

Native speakers of English reject the corresponding examples with *but* as ungrammatical.

(150) a. *Most girls but Ann passed the test.
b. *Many girls but Ann passed the test.

These examples contrast with the perfectly acceptable cases reported in Peters and Westerståhl (2023), where *but* is compatible with *most* and *many*, contributing a negative inference in such contexts, like the one in (151). Again, the reported judgment for such cases is that including Sundays does not necessarily make the *most*-claim false.

(151) We are available most days but Sundays.

We remain optimistic that a compositional theory in terms of *EXH* incorporating the

key insights from Peters and Westerståhl (2023) can be developed to account for all the contrasts observed here. For now, we must leave this to future research.

There is also the issue of the difference in the ability of *besides* and *but* to occur with the definite article. While the former is acceptable with both singular and plural on the noun and gives rise to an exceptive inference, the latter is not.

(152) a. The student(s) besides Ann passed. ⇝ Ann didn't pass
b. *The student(s) but Ann passed.

(152a) is arguably in line with our predictions. It means that the maximal student entity who does not overlap with Ann passed and that the maximal student plurality including Ann did not pass. In the case of singular, maximality in the first conjunct reduces to uniqueness. Recall that we adopt the view that pluralities are present in the alternatives used by *EXH* as here the singular feature on the noun can be ignored. If that is on the right track, one question that will have to be addressed relates to homogeneity. Plural definites with distributive predicates give rise to a homogeneity inference. When negated, all atomic parts of the plurality should make the predicate false. In the case of (152a), however, this would lead to a clash with the requirements of the prejacent.

The case of (152b) is more problematic. With singular we predict it to mean that the unique student who is not Ann passed and that the unique student who is not the unique student who is not Ann did not pass. In other words, there are two students one of who is Ann and only the one who is not Ann passed. For (152b) with plural predictions fully parallel to those with universal quantifiers are made, modulo the issue of homogeneity mentioned. That is, (152a) and (152b) are both predicted to be acceptable, again ignoring the issue of homogeneity, and should be equivalent. One way to reconcile the predictions for (152a) with the unacceptability for (152b) would be to posit a lexical difference between *besides* and *but*. In particular, the latter might be specified so as to take a quantificational determiner as one of its arguments, as in (153), following von Fintel (1994), whereas the former could stay as proposed above.

(153) $[\![but]\!] = \lambda X_e.\lambda f_{<et>} : \forall z[z \leq X \rightarrow f(z) = T].\lambda Q_{<<et><et,t>>}.$
$Q(\lambda Y.f(Y) = T \,\&\, Y \oslash X)$

With *the* not being a quantificational determiner, (152b) would result in undefinedness. This strategy would raise the question why definite determiners could not be shifted to the type of quantificational determiners to make (152b) acceptable.

Finally, the difference between (154a) and (154b) is also left unaccounted for at the moment. Again (154a) is in line with our predictions. It would say that the number of students who passed without Ann being taken into account is zero but not so with Ann considered. From this it correctly follows that Ann passed. The same prediction is, however, made for (154b). It is not obvious that the suggested modification for *but* in (153) would remedy this situation. Even on Bylinina and Nouwen's 2018 account making reference to a null individual, *zero* in combination with *many* yields a quantificational determiner denotation suitable for (153). What is more, their obligatory exact-interpretation for *zero* would allow for exceptive *but* to be licensed through *EXH*.

(154) a. Zero students besides Ann passed ⇝ Ann passed
b. *Zero students but Ann passed

References

Bar-Lev, M. E. and D. Fox (2020). Free choice, simplification, and innocent inclusion. *Natural Language Semantics 28*, 175–223.

Beck, S. and H. Rullmann (1999). A flexible approach to exhaustivity in questions. *Natural Language Semantics 7*(3), 249–298.

Ben-Avi, G. and Y. Winter (2003). Monotonicity and collective quantification. *Journal of Logic, Language and Information 2*(12), 127–151.

Buccola, B. (2015). Variable monotonicity and *less than*: when van benthem's problem is not a problem. In *Proceedings of the 45th annual meeting of the North East Linguistic Society (NELS 45)*, Volume 1, pp. 125–134. UMass, Amherst.

Buccola, B. (2016). Severing maximality from fewer than: evidence from genericity. In *Proceedings of Sinn und Bedeutung 19*, pp. 127–144. University of Konstanz.

Buccola, B. and B. Spector (2016). Modified numerals and maximality. *Linguistics and Philosophy 39*(3), 151–199.

Bylinina, L. and R. Nouwen (2018). On "zero" and semantic plurality. *Glossa: A Journal of General Linguistics 3*(1), 1–23.

Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford: Oxford University Press.

Chierchia, G., D. Fox, and B. Spector (2012). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, and K. von Heusinger (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 3. Berlin: Mouton de Gruyter.

Cresti, D. (1995). Extraction and reconstruction. *Natural Language Semantics 3*(1), 79–122.

Crnič, L. (2018). A note on connected exceptives and approximatives. *Journal of Semantics 35*(4), 741–756.

Crnič, L. (2021). Exceptives and exhaustification. In *Proceedings of WCCFL 39*.

Crnič, L. (2024). A distributed analysis of *only*. *Linguistics and philosophy 47*, 611–652.

Dayal, V. (1996). *Locality in Wh-Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer Academic Press.

Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry 30*(2), 157–196.

Fox, D. (2002). Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry 33*(1), 63–96.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. Basingstoke, New York: Palgrave.

Fox, D. and M. Hackl (2006). The universal density of measurement. *Linguistics and Philosophy 29*, 537–586.

Fox, D. and R. Katzir (2011). On the characterization of alternatives. *Natural Language Semantics 19*, 87–107.

Fox, D. and B. Spector (2018). Economy and embedded exhaustification. *Natural Language Semantics 26*(1), 1–50.

Gajewski, J. (2002). L-analyticity in natural language. ms. MIT.

Gajewski, J. (2008). NPI *any* and connected exceptive phrases. *Natural Language Semantics 16*, 69–110.

Gajewski, J. (2011). Licensing strong npis. *Natural Language Semantics 19*(2), 109–148.

Gajewski, J. (2013). An analogy between a connected exceptive phrase and polarity items. In E. Csipak, R. Eckardt, M. Liu, and M. Sailer (Eds.), *Beyond 'Any' and 'Ever': New Explorations in Negative Polarity Sensitivity*, Boston, pp. 183–212. de Gruyter.

García Álvarez, I. (2008). *Generality and exception: A study in the semantics of exceptives*. Ph. D. thesis, Stanford University, Stanford.

Groenendijk, J. and M. Stokhof (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. Ph. D. thesis, University of Amsterdam, Amsterdam.

Hackl, M. (2000). *Comparative Quantifiers*. Ph. D. thesis, Massachusetts Institute of Technology.

Hamblin, C. (1973). Questions in Montague grammar. *Foundations of Language 10*(1), 41–53.

Heim, I. (1991). Artikel und definitheit. In A. von Stechow and D. Wunderlich (Eds.), *Semantik: Ein Internationales Handbuch der Zeitgenössischen Forschung*, pp. 487–535. Berlin: Walter de Gruyter.

Heim, I. (2000a). Degree operators and scope. In B. Jackson and T. Matthews (Eds.), *Proceedings of SALT X*, Cornell University, Ithaca, NY, pp. 40–64. CLC Publications.

Heim, I. (2000b). Notes on interrogative semantics. Lecture notes for Advanced Semantics, MIT.

Heim, I. (2008). Features on bound pronouns. In D. Harbour, D. Adger, and S. Béjar (Eds.), *Phi-Theory: Phi-Features across Modules and Interfaces*, pp. 35–56. Oxford, New York: Oxford University Press.

Heim, I. (2019). Functional readings without type-shifted noun phrases. In M. Krifka and M. Schenner (Eds.), *Reconstruction Effects in Relative Clauses*, pp. 283–302. Berlin: De Gruyter.

Hirsch, A. (2016). An unexceptional semantics for expressions of exception. In *University of Pennsylvania Working Papers in Linguistics*, Volume 22.

Hoeksema, J. (1987). The logic of exception. In *Proceedings of the Fourth Eastern States Conference on Linguistics*, Columbus, OH, pp. 100–113. The Ohio State University.

Horn, L. and S. Bayer (1984). Short-circuited implicature: A negative contribution. *Linguistics and Philosophy 7*, 397–414.

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

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy 1*(1), 3–44.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy 30*, 669–690.

Landman, F. (2011). Boolean pragmatics. In J. van der Does and C. Dulith Novaes (Eds.), *This is not a festschrift*. Universiteit van Amsterdam.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*. Berlin, New York: de Gruyter.

Liu, M. (2019). Explaining the exceptive-additive ambiguity in mandarin. In *Proceedings of the 22nd Amsterdam Colloquium*, pp. 475–483. Amsterdam: ILLC, UvA.

Longobardi, G. (1987). Extraction from NP and the proper notion of head government. In A. Giorgi and G. Longobardi (Eds.), *The syntax of noun phrases*, pp. 57–112. Cambridge: Cambridge University Press.

Mayer, R. (1993). Domain restriction and other kinds of reference set operations. *Theoretical Linguistics 19*, 129–200.

Mayr, C. and M.-C. Meyer (2014, September). More than at least. talk at Two Days at least workshop.

Mayr, C. and E. Vostrikova (2023). Everyone except possibly ann. In *Semantics and Linguistic Theory (SALT) 33*, pp. 131–150.

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

Nadathur, P. and D. Lassiter (2018). An experimental look at the negative implications of exceptives. In S. Hucklebridge and M. Nelson (Eds.), *Proceedings of NELS 48*, Boston, pp. 201–210. Amherst: GLSA.

Nouwen, R. (2008). Upper-bound *no more*: the exhaustive interpretation of non-strict comparison. *Natural Language Semantics 16*(4), 271–295.

Oskolskaya, S. (2011). Constructions with the meaning of exception of a participant in Russian. Master's thesis, St.Petersburg State University.

Peters, S. and D. Westerståhl (2006). *Quantifiers in language and logic*. Oxford: Oxford University Press.

Peters, S. and D. Westerståhl (2023). The semantics of exceptives. *Linguistics and Philosophy 46*, 197–235.

Rullmann, H. (1995). *Maximality in the semantics of Wh-Constructions*. Ph. D. thesis, University of Massachusetts, Amherst.

Rullmann, H. and S. Beck (1998). Presupposition projection and the interpretation of *which*-questions. In D. Strolovitch and A. Lawson (Eds.), *Proceedings of SALT VIII*, Ithaca, NY, pp. 215–232. CLC Publications.

Sauerland, U. (2003). A new semantics for number. In R. B. Young and Y. Zhou (Eds.), *Semantics and Linguistic Theory (SALT) 13*, Ithaca, NY, pp. 258–275. CLC Publications.

Sauerland, U. (2013). Presuppositions and the alternative tier. In *Proceedings of SALT 23*.

Schlenker, P. (2012). Informativity-based maximality conditions. *Snippets 26*, 18–19.

Sevi, A. (2008). A two-layered analysis of the hebrew exceptive *xuc mi*. In S. Armon-Lotem, G. Danon, and S. Rothstein (Eds.), *Current Issues in Generative Hebrew Linguistics*, pp. 337–352. John Benjamins Publishing Company.

Spector, B. (2013). Global positive polarity items and obligatory exhaustivity. *Semantics and Pragmatics 7*(11), 1–61.

Spector, B. (2014). Plural indefinites and maximality. UCLA colloquium talk.

von Fintel, K. (1989). Exception phrases. In *University of Massachusetts Occasional Papers in Linguistics*, Volume 15, pp. 1–8.

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

von Fintel, K. (1994). *Restrictions on Quantifier Domains*. Ph. D. thesis, University of Massachusetts-Amherst.

von Fintel, K., D. Fox, and S. Iatridou (2014). Definiteness as maximal informativeness. In L. Crnič and U. Sauerland (Eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*, Volume 1, pp. 165–174. Dordrecht: Foris Publications.

Vostrikova, E. (2019a). Compositional analysis for clausal exceptives. In K. Blake, F. Davis, K. Lamp, and J. Rhyne (Eds.), *Semantics and Linguistic Theory (SALT) 29*, Ithaca, NY, pp. 420–440. CLC Publications.

Vostrikova, E. (2019b). *Phrasal and clausal exceptive-additive constructions crosslinguistically*. Ph. D. thesis, UMass Amherst.

Vostrikova, E. (2021). Conditional analysis of clausal exceptives. *Natural Language Semantics 29*, 159–227.

Winter, Y. (2001). *Flexibility principles in boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge, MA: The MIT Press.

Zuber, R. (1998). On the semantics of exclusion and inclusion phrases. In A. Lawson (Ed.), *Proceedings of SALT 8*, pp. 267–283. Cornell University Press.