Even and _Only_: Arguing for parallels in scalarity and in constructing alternatives[*]
(To appear in the proceedings of NELS49)

Yael Greenberg

Bar Ilan University

Abstract: This paper examines three types of entries assigned to the particle 'only' in the literature, namely 'non–scalar', 'scalar' and 'hybrid' entries, and raises novel arguments favoring the 'hybrid' entry over the more popular 'non-scalar' and 'scalar' ones. We propose that adopting this entry allows us to capture the long observed 'scalar antonymy' between 'only' and 'even' – where both particles have scalar presuppositions with universal force, requiring p to be weaker (in the case of only) / stronger (in the case of even) than all other alternatives in C. Adopting this entry is also supported by that fact that it does a better than the more popular 'non-scalar' and 'scalar' entries in explaining the infelicity of 'only' where sentences weaker than p are salient in the discourse, and the fact that this infelicity is the mirror image of infelicity patterns with 'even'. Importantly, however, unlike the 'scalar ordering' antonymy, we argue that 'only' and 'even' should not be considered 'mirative' /'evaluative' antonyms: While 'even' is indeed a real evaluative particle, presupposing that p (as well as its alternatives) indicates a degree which is higher than the salient standard on the relevant scale, apparent mirror-imaged evaluativity inferences of only (indicating a degree below expectations / the salient standard) should not hardwired into its entry, but are rather pragmatically derived from the universal quantification over alternatives and constraints on constructing C. Finally, we show that maintaining the universal quantification over alternatives in C in the scalar presuppositions of both 'only' and 'even' faces parallel challenges, and discuss some ways to handle them.

1. Introduction

An observation found in the literature about the focus sensitive particles _only_ and _even_ is that the two express some sort of opposite or mirror imaged scalar relation. Zeevat 2008, for example, argues that "_only_...expresses that the size of something is disappointingly small: one expected more. Similarly, _even_ expresses that one expected less. (p. 301).

Beaver & Clark 2008 cite (1) and write that "whereas (1a) is appropriate if wearing a bow tie when teaching is **less**, e.g., eccentric than had been expected or previously indicated, (1b) is appropriate if wearing a bow tie when teaching is regarded as significantly **more**, e.g., eccentric than has been expected or previously indicated" (p. 71):

(1)   a.   David only wears a bow tie when [teaching]$_F$
        b.   David even wears a bow tie when [teaching]$_F$

However, the entries often assigned to *only* and *even* in the literature do not capture this observation. *Even* is usually characterized as in (2), presupposing that the prejacent, p is the strongest alternative in C (the set of contextually relevant focus alternatives) on a scale, and to assert p (cf. Horn 1969, Karttunen & Peters 1979, Rooth 1992).[1] As to *only*, we will consider here three types of entries assigned to it in the literature. Two rather popular entries are the **non-scalar** and the **scalar** entries in (3a)-(3b), asserting that all distinct / all stronger alternatives to p in are false, respectively. The third is what we will call a **hybrid** entry in (3c), including both a scalar presupposition, requiring p to be the weakest alternative in C, and a non-scalar assertion, negating all alternatives distinct from p in C:[2]

(2) $\|\mathbf{\mathit{even}}\|^{g,c} = \lambda C. \lambda p. \lambda w: \forall q \in C\ q \neq p \rightarrow p >_C q.\ p(w) = 1$

(3)   a. $\|\mathbf{\mathit{only}}_{non\text{-}scalar}\|^{g,c} = \lambda C.\lambda p.\lambda w:p(w)=1.\forall q \in C\ q \neq p \rightarrow q(w)=0$
        b. $\|\mathbf{\mathit{only}}_{scalar}\|^{g,c} = \lambda C.\lambda p.\lambda w:p(w)=1.\forall q \in C[q \neq p \wedge q >_C p] \rightarrow q(w)=0$
        c. $\|\mathbf{\mathit{only}}_{hybrid}\|^{g,c} = \lambda C.\lambda p.\lambda w:p(w)=1 \wedge \forall q \in C\ q \neq p \rightarrow p <_C q. \forall q \in C$
                $q \neq p \rightarrow q(w) = 0$

As its name suggests, the **non-scalar** entry of *only* (cf. Horn 1969, Karttunen & Peters 1979, Rooth 1992) does not contain any reference to scales, and hence does not capture the observation of the scalar antonymy with *even*. In contrast, the **scalar** entry clearly makes reference to scales, namely to either non-entailment /'rank order' scales (as in *John is only a child*, negating, e.g. *John is a grown-up*) or to plurality-based / entailment scales, (as in *Only John arrived*, negating e.g. *John and Mary arrived*) (cf. Klinedinst 2005, Beaver & Clark 2008, Roberts 2011, Alxatib 2013, Coppock & Beaver 2014, Chierchia 2013). Despite this, this entry is also unable to capture the observation about the scalar antonymy with *even*: Unlike the scalar presupposition of *even* in (2) which dictates that p is stronger than all its alternatives in C, the scalar assertion of *only* in (3b) does not dictate any such relation at all. All it requires is that those alternatives in C which are stronger than p on the scale are false.

In contrast to these two entries, the **hybrid** entry in (3c) (cf. Guerzoni 2003, inspired by König 1991), gets closer to capturing the scalar 'mirror imaged' relation. In particular, the scalar presupposition in this entry, requiring p to be the weakest alternative in C, is the

---

[1] As is well known, there are several debated components in this entry of *even*. These concern, for example, the presence / absence of an additive presupposition in it, the universal quantification in the scalar presupposition and the nature of the scale. We will ignore in this paper the debate about the additive presupposition, and will deal with the two other debated components below.

[2] We are ignoring here many subversions of these entries, and adopting the debated assumption that the prejacent of *only*, p, is presupposed (Horn 1969, Rooth 1992).

mirror image of the scalar presupposition of *even,* requiring *p* be the strongest alternative in C. If indeed one accepts the observation that *only* and *even* are scalar opposites, this seems to be an advantage of this entry over the **non-scalar** and **scalar** entries.

The goal of this paper is to provide arguments for the **hybrid** entry of *only* over the **non-scalar** and **scalar** ones, which although still concern its relationship with *even* are more empirical in nature. To do that we examine in section 1 a challenge for the **scalar** and **non-scalar** entries, namely the infelicity of *only* in cases where weaker material than the one expressed in its prejacent, *p,* is salient in the discourse. In section 2 we argue that attempting to address this challenge by hardwiring into these two entries an 'evaluative' presupposition is problematic. Section 3 argues that the **hybrid** entry provides a general solution for to the infelicity puzzle and captures the fact that *even* exhibits a mirror imaged infelicity pattern. Section 4 examines challenges for the universal quantification over alternatives in C in the semantics of both *even*, given the entry in (2), and *only,* given the **hybrid** entry in (3c). In section 5 we suggest that the evaluativity effects of *only* can be derived from the interaction between its **hybrid** entry and constraints on alternatives in C (following ideas in Krifka 2000). Section 6 concludes.

1.   An 'infelicity puzzle' for the non-scalar and scalar entries of *only*

Orenstein & Greenberg 2013, Orenstein 2016 examines the infelicity of *only* in (4a), with a numeral scale, where material weaker than the parallel material in the prejacent is salient. (4b-c) illustrate a similar phenomenon with plurality and rank-order scales:

(4)   a.   Mira has 2 kids, Paul has 4 kids, but Jim has (#*only) 3* kids
       b.   Henry won the gold medal, Bill won bronze and John (#only) won [silver]$_F$
       c.   Two years ago John interviewed the prime-minister. Last year he interviewed the minister's assistant and this year he (#only) managed to interview [the minister]$_F$

We follow here the assumption (Rooth 1992, Fox & Katzir 2011, Katzir 2014), that salient material is used to construct alternatives in C, leading to the sets in (5a-c):

(5)   a.   {…Jim has 2 kids, Jim has 4 kids, <u>Jim has 3 kids</u>}
       b.   {…John won gold, John won bronze, <u>John won silver</u>}
       c.   {… John interviewed the prime minister John interviewed the minister's assistant, <u>John interviewed the minister</u>}

Given such C sets, neither the **non-scalar,** nor the **scalar** entries of *only* can explain the infelicity of this particle in (4)) in a simple way. While they require the negation of all alternatives distinct from *p* in C (given the **non-scalar** entry), or those stronger than *p* in C (given the **scalar** entry), neither of them rule out the presence of some weaker alternatives in C. Using just these entries, then, the infelicity of *only* in (4) remains a puzzle.[3]

---

[3] Notice that if, e.g. (4a) is replaced by (ia), and if we make the further assumption that C contains just *p* and alternatives constructed based on discourse salient material, as in (iib), then the infelicity of *only* could be in principle attributed to a constraint against non-vacuity of *only* (cf. Crnič 2011, Alxatib 2013):

2. An attempt: Using an 'evaluativity' constraint on *only* for solving the infelicity puzzle

At this point one can try to explain the infelicity of *only* in (4a-c) by augmenting the **scalar** or **non-scalar** entries in (3a-b) with a 'mirative' constraint which requires that *p* indicates a degree or a measure which is lower than what is expected (Zeevat 2008), or alternatively which expresses 'not so much / not a lot', or which is lower than most / sufficiently many alternatives (Klinedinst 2005). A more general characterization of this constraint (Alxatib 2013) requires *p* to be 'evaluative', i.e. to indicate a degree lower than a salient norm or standard on a relevant scale similar to the norm or standard referred to by covert degree modifier *POS* with gradable adjectives. Alxatib 2013 convincingly emphasizes that this norm or standard need not correlate with what is expected. Following this latter formulation, we will henceforth refer to such a constraint as an 'evaluativity constraint'. Imposing such a constraint on *only* has two types of motivations in the literature. The first is the need to explain "smallness" inferences of sentences like (6):

(6)   John only wrote 2 papers (writing 2 papers is not a lot / less than what is needed)

The second motivation is to explain infelicitous uses of *only* where such inferences seem to clash with other material or with real world knowledge, like (7a-c):

(7)   a.   I really expected a single room with 2 beds but (#only) got [a suite.]$_F$
               (Beaver & Clark 2008 p.252)
        b.   The average score on the exam was a C. Mary (#only) got an [A-]$_F$
               (Klinedinst 2004 p. 4)
        c.   John (#only) has [8]$_F$ kids (uttered in typical western contexts where 8 kids is a lot)

Now, one can suggest that perhaps *only* is infelicitous in sentences like (4) above because, given the 'weaker' sentence in the context in all of them *p* indicates 'a lot', 'more than expected', or 'more than the standard', etc., similarly to what we find in (7).

     However, such an attempt faces two main issues. First, evaluativity (indicating 'not a lot') does not seem to be a necessary condition for the felicity of *only*. Second, it is not a sufficient condition either. Let us look at each of these issues in terms.

     Doubts regarding the claim that expressing 'less than expected' is a necessary condition on the felicity of *only p* were raised in e.g. Roberts 2011, Orenstein & Greenberg 2013, Orenstein 2016 and others. Consider (8):

(8)   (Context: My mother and I are organizing a weekend for the whole family. We discuss the location where everyone will stay. Mom says:) Rina has four kids so

---

(i)   a.   Mira has 2 kids, Paul has 4 kids, but Jim has (#*only*) *3* kids
        b.   {Jim has 2 kids, <u>Jim has 3 kids</u>}

That is, in this case we could claim that *only* is infelicitous in (ia) since there are no stronger alternatives in C to negate. Such a claim, though, cannot be made for the (4a) above: As can be seen in (5a), here there ARE stronger alternatives to reject, so *only* does not operate vacuously. Nonetheless it is infelicitous.

she'll stay in this apartment. Esti has only three kids, so she can stay in that apartment. (Orenstein & Greenberg 2013 #(9)).

As Orenstein & Greenberg 2013, Orenstein 2016 point out, in (8) *only* is felicitous although here *p* does not seem to be considered less than expected. Instead the use of *only* here seems to be licensed simply due to the comparison with the higher number of children of Rina.

A similar point can be seen if we take a more general characterization of the constraint on *only,* as requiring *p* to be below a salient standard on the scale:

(9)   a.   (The average score on the exam is a B / To get accepted one needs at least a B)
                John: Wow! I got an A+!
                Bill: You got (even) better than me, then! I only got an A. I am so happy!
        b.   (The average height for men here is 1.78m).
                A: John is tall. He is 16 years old and already 1.85m tall
                B: (a 15 year old boy): He is (even) taller than me, then – I am only 1.83m

Given the context in (9a), getting an A is more than the salient standard (calculated based on the distribution of grades, or on what is needed to get accepted). Nonetheless *only* is felicitous. Moreover, *only* keeps being felicitous not only if Bill is a wonderful student (so getting an A might be considered below his individual standard), but also if both he and John are average students, or if he is an average student and John is a great student, i.e. even if getting an A is clearly above his standard of success or what Bill expects to get. A similar picture is seen in (9b), where B's degree of tallness is higher than the standard given the average population, or B's age group. Nonetheless *only* is felicitous as well. Again, what seems to license *only* in both cases is not its 'evaluative' nature, i.e. the fact that *p* indicates a degree which is less than the relevant standard, but its 'comparative' nature, i.e. that *p* indicates less than degrees in the sentences before (i.e. John being 1.85m tall, John's getting an A+), crucially even though this degree does not constitute a 'standard' at all.
An interesting support of this conclusion is the observation that, as seen in (9a,b), *only* continues to be felicitous when we add *even* to the sentences before. This is significant since it was independently claimed (cf. Greenberg 2015, 2018) that *even* is not only 'comparative', but also necessarily 'evaluative', indicating '**more** than the standard' for both *p* and its alternatives in C. This is motivated by inferential patterns in e.g. (10b), where it is only the presence of *even* which entails that both John and Bill are tall (cf. 10a). Notice also that unlike the cancellable evaluativity inferences of *only* in (9), the parallel ones with *even* in (10) are non-cancellable. As the infelicity of *even* in (11a,b) show, for this particle the 'comparative' requirement that *p* is higher than a salient degree is not enough:

(10)   a.   John is 1.70m tall, and Bill is 1.75m / taller (both can be short).
        b.   John is 1.70m tall and Bill is even 1.75m / taller (neither can be short)

(11)   a.   (The average score on the exam is a B / The minimum grade for acceptance is a B)
                John: Well, I got a C in the exam…
                Bill: and I #(even) got a C+

b.   (The average height for men here is 1.78m).
A: John is short. He is 1.60m.
B: And Bill is (#even) 1.65m.

To capture such effects, Greenberg 2015, 2018 suggests a 'gradability-based' formulation of the scalar presupposition of *even*. This has both a 'comparative' scalar ordering component, requiring *p* to be 'stronger' than its all alternatives *q* in C (that is, indicating a higher degree on the contextually supplied scale), as well as a hardwired 'evaluative' component, requiring both *p* and its alternatives to indicate a degree which is above the contextually supplied standard on that scale.[4] For example in (10b), *Bill is even 175m. tall* presupposes (a) that Bill's degree of tallness (or suitability for the basketball team, etc.) is higher in the accessible worlds where he is 1.75m tall than in those where is 1.70m tall and not 1.75m and (b) that in the latter worlds he is at least as tall as the standard, i.e. that he is considered POS tall when he is 1.70m tall (and of course also when he is 1.75m).

Let us turn back now the interaction between *only* and *even* in (9) above. In (9a), the presence of *even* in Bill's first sentence (*You even got higher than me, then*) presupposes that that both getting an A+ and getting an A are at least as high as the standard on the relevant scale, e.g. of success, i.e. both are considered POS successful. Crucially, then, if *only* indeed presupposed that its prejacent indicates 'less than the standard' we would wrongly predict it to be infelicitous in (24) due to a clash with the opposite presupposition of *even*.[5] A similar argument can be for (9b). More generally, this data seems to show that evaluativity (in this case presupposing 'less than the standard') should not be hardwired into the semantics of *only*, as it is not a necessary condition for its felicity.[6]

In addition, evaluativity does not seem to be a sufficient condition on the felicity of *only* either. Consider (12):

(12)   (Context: How many papers did your faculty members write during the last 3 years?) Let's see: Susan and Ann wrote 10 papers, Sam wrote 8, Henry wrote 5, Tom wrote 7, Ted wrote 6, Ian wrote 3, and Bill (#only) wrote 4.

This is, in fact, another illustration of the 'infelicity puzzle' of *only,* discussed in section 1. What is important here is that the prejacent *Bill wrote 5 papers* indicates a quantity which is clearly lower than most alternatives based on the previous sentences, and it can easily count as 'less than expected / less than the standard', etc. Nonetheless *only* is infelicitous. Intuitively, this is due to the presence of the weaker sentence *Ian wrote 3* in the context.

To conclude, hardwiring evaluativity into the **scalar** or **non-scalar** entries of *only* seems neither independently justified, nor sufficient to help solve the 'infelicity puzzle'.

3.   Solving the infelicity puzzle using the 'hybrid' entry of *only*

---

[4] See Greenberg 2016, 2018 for arguments against the comparative likelihood formulation of the scalar presupposition of *even,* and a formulation of a 'gradability-based' presupposition. See also Greenberg (to appear) for an analysis of the Hebrew *even*-like particle *bixlal*, further supporting this analysis.
[5] One could try to argue that two different standards are triggered by *even* and by *only* in (9a) (and in (9b)). However, this has no independent motivation. See also Alxatib 2013, arguing that similar 'standard shifts' are generally banned with sequences of expressions with opposite scalar inferences (like *already* and *late*).
[6] This still leave the question of how the evaluativity effects of *only*, seen in e.g. (6)-(7), should be explained. We turn to this question in section 5.

How, then, can the 'infelicity puzzle' of *only* be explained? Orenstein & Greenberg 2013, Ornstein 2016 suggest to augment the **scalar** entry of *even* (as in (3b) above) with an explicit presupposition requiring that all 'salient' alternatives in C (i.e. those constructed based on salient material) to be stronger than p:

(13) $\lambda C.\lambda p.\lambda w: w \in p \wedge \forall q \ [q \in C \wedge q \text{ is salient} \wedge q \neq p] \rightarrow q >_s p. \ \forall q \ [[q \in C \wedge q >_s p] \rightarrow w \notin q]$

While this move immediately derives the infelicity of *only* in (4) and (12) above, we will take a slightly different path in trying to solve the puzzle. This has two main reasons. First, the fact that (13) makes different requirements on different sets of alternatives, namely an assertion concerning all alternatives in C and a presupposition only over a 'salient' subset of C, is unusual. Focus sensitive particles usually dictate a relation between p and a constrained set of alternatives, and not different relations on different sets.

More importantly, the entry in (13) does not capture the fact that the infelicity of *only* seems to be part of a more general pattern, as it has a parallel with *even*. To see that consider (14)-(15), where the (in)felicity of *even* is a mirror image of *only*:

(14)   a.   Bill wrote 5 papers this year. John (#only /even) wrote [6]$_F$
       b.   Bill won the bronze medal. John (#only / even) won the [silver medal]$_F$

(15)   a.   Bill wrote 6 papers this year. John (only / #even) wrote [5]$_F$
       b.   Bill won the silver medal. John (only / #even) won the [bronze medal]$_F$

Remember also that in (12) above we saw that for *only* to become infelicitous, it is enough that there is one sentence weaker than p in the context, crucially, even if all other sentences are stronger than p. We can now see the precise mirror image of this situation in (16), where there is one salient alternative stronger than the prejacent, and *even* is infelicitous:

(16)   (Context: How many papers did your faculty members write during the last 3 years?) Let's see: Susan and Ann wrote 5 papers, Henry wrote 12, Tom wrote 10, Ted wrote 9, Sara wrote 7, Ann wrote 15, Ian wrote 13, and Bill (#even) wrote 14.

The infelicity of *only* in the presence of sentences weaker than p in the context, then, seems is a mirror image of the infelicity of *even* in the presence of sentences stronger than p.

But crucially, unlike this infelicity of *only,* which as we saw above constitutes a puzzle and cannot be derived using commonly-used **non-scalar** and **scalar** entries, this infelicity of *even* is not puzzling at all, as it is easily derivable from its traditional lexical entry in (2) above. In particular, assuming again that discourse salient material is used to construct alternatives to p in C (Fox & Katzir 2011, Katzir 2014), the infelicity of *even* in (15)-(16) is straightforwardly explained, since in each of these sentences the C set has an alternative stronger than the prejacent, constructed based on salient material. This leads to the failure of the scalar presupposition of *even,* requiring p to be the strongest alternative in C.

This now naturally leads the way to constructing an entry of *only* with a parallel component, i.e. which presupposes that p is the weakest alternative in C. The Guerzoni

2003 style **'hybrid'** entry in (3c) above fulfills this condition.[7] Given this entry, the infelicity of *only* discussed above is no longer a puzzle: It is straightforwardly derived in a parallel fashion to the derivation of the infelicity of *even*: Here too the scalar presupposition fails, this time due to the presence of a salient alternative in C weaker than *p*.

4. Issues for the universal quantification over alternatives with *only* and *even*

A central feature of the **hybrid** entry of *only*, then, is that its scalar presupposition involves universal quantification over the alternatives to *p* in C: ALL such alternatives must be stronger than *p*. This feature was motivated by the infelicity of *only* in sentences like (4) and (12). However, it is risked by the felicity of *only* in sentences like (17a,b), where C can arguably contain propositions like *Bill won the bronze medal* in (17a), and *Bill solved 6 problems* in (17b), which are weaker than *p*, and can easily count as relevant alternatives:

(17)   a.   John won the gold medal. Bill only won silver
        b.   John solved 8 questions in the exam. Bill only solved 7

Importantly, similar kinds of doubts were independently raised against the universal quantification over alternatives in the scalar presupposition of *even*, requiring *p* to be stronger than ALL alternatives in C. A well-known example illustrating these doubts is the felicity of *even* in (18), where there is an alternative stronger than *p*, i.e. *Mary made it to the finals*, can be easily considered 'contextually relevant', and hence part of C:

(18)   Not only did Mary win her first round match, she even made it to [the semifinals]F (Kay 1990 (p. 89))

Kay 1990 took such examples to argue that *even* does not presuppose *p* to be stronger than all alternatives in C, but stronger than a 'context proposition'. Similar challenges to the 'universal' semantics of *even* are illustrated in (19), where *Bill won gold* and *Bill solved 10 problems* can be arguably considered contextually relevant and hence members of C:

(19)   a.   John won the bronze medal. Bill even won silver
        b.   John solved 8 problems in the exam. Bill even solved 9

Notice, again, the striking parallel between *only* and *even*. It seems that such examples seem to pose two main options for BOTH particles. One is to keep universal quantification and find a way to restrict C so that the felicity of *even / only* in cases like (17)-(19) is derived. Another is to abandon universal quantification over alternatives (again, for both *even* and *only*), and find a way to derive the infelicity of these particles in e.g. (4) and (12) for *only,* and (15)-(16) for *even*. In what follow we examine these two options.
On the first option we keep universal quantification over alternatives in C for both *only* and *even*, and find a way to appropriately restrict C. For example, we can assume that discourse salient material has a priority over the lexicon in constructing alternatives in C, so when alternatives based on discourse salient material (e.g. *John won bronze* in (19a)) are presen,t those based on lexicon (*John won gold*) can be left out of *C*. This kind of

---

[7] Cf. Crnič 2012, Charnavel 2017, Grubic 2015, Liu 2017 (for the Chinese exclusive *jiu*).

suggestion can be seen as refining the algorithm for constructing alternatives in e.g. Fox & Katzir's 2011, Katzir's 2014, which originally takes both the lexicon and salient constituents to be sources for substituting material for constructing alternatives.[8]

This option is supported for *even* by a variant of Kay's felicitous example in (18), namely the infelicitous (20), cited in Greenberg 2016. Unlike (18) here the stronger alternative is constructed based on discourse salient material (the explicit "Harry made it to finals") and not based on the lexicon. Greenberg 2016 suggests that *even* is infelicitous in this case since, being based on discourse salient material this stronger alternative cannot be ignored anymore and it must enter C, so the universal scalar presupposition necessarily fails. The felicity contrasts in (21) and (22) now further support this idea, for both *even* and *only*:

(20)   (Harry, John and Bill participated in the sports competition.) Harry made it to the finals, John won his first round match, and Bill (??**even**) made it to the [SEMI-finals]F. (Greenberg 2016, # (7)).

(21)   a.   Bill won bronze in the competition. John even won silver
        b.   Henry won gold. Bill won bronze, and John (#even) won silver

(22)   a.   Bill won gold in the competition. John only won silver
        b.   Henry won bronze. Bill won gold and John (#only) won silver

However, a challenge for this first option are cases like (23)-(24). (23a), pointed out in Greenberg 2019, involves 'sequences of *only*'. In particular, *Bill only has 3* is felicitous in (23a) although there is a discourse salient weaker sentence *Bill has 2*, which is itself modified by *only*. Xiang 2019 makes a similar observation for *only* in (23b) and points out a parallel case with sequences of *even*, as in (24) (a variant of the infelicitous (20) above):

(23)   a.   The average number of children here is 5. John **only** has 2 children. And Bill **only** has 3. (Greenberg 2019, footnote #12)
        b.   [— How much are these shoes? —
                Well, this pair is **only** $40, and that pair is (**only**) $50. (Xiang 2018 p. 27)

(24)   [— Harry, John and Bill participated in the sports competition. I heard that Harry won his first round. How exciting! — Well,] not only that Harry won his first round, John **even** made it to the finalsF, and Bill also **even** made it to the semi-finalsF! (Xiang 2018 p. 26)

Xiang takes the felicity of *even* in (24) to support the second option mentioned above, namely abandoning the universal quantification over alternatives in C, and adopting an existential scalar presupposition (requiring p to be stronger than at least one alternative in C). In principle, a similar option can be supported for *only* as well.[9]

This second option, though, should still explain the original infelicity of *even* and *only* described in above (e.g. in section 1 and 3). Xiang 2019 suggests an interesting explanation of this infelicity, and consequently, the generalization in (25):

> Contra Greenberg (2016, 2019), I argue that the oddness of *even* in [(20)] isn't due to a failure of satisfying the scalar presupposition of *even*. Instead, it is due to the oddness of not using *even* when the option of using *even* is clearly available… a conjunction of the form "S1 and *even*-S2" implicates that *even*-S1 is infelicitous (or more precisely, that the evaluative scalar presupposition, that S1 is unlikely, is false)…in contrast to [(20)], *even* felicitously appears in the *semi-finals*-clause in [(24)] as it also appears in the *finals*-clause. (Xiang 2019, p. 26)

(25) For an evaluative expression δ, a coordination with clauses {p, δ(q)} is felicitous only if the evaluative inference of δ(p) is false. (Xiang 2019, p. 26)

This explanation seems to predict that a sentence with *only p / even p* will be infelicitous if another sentence prior to it in the coordination[10] which can in principle be modified by *only / even* (i.e. for which the evaluative inference is true) remains unmodified. However, this predication does not seem to be borne out. Let us show that for *only*:

(26) A: How many papers did your faculty members write in this period?
B: Well, John did great. He wrote 8 papers. The rest didn't do so well: Bill (only) wrote 5 and Susan only wrote 4.
B': Well, John did great. He wrote 8 papers, The rest didn't do so well: Bill (only) wrote 5, Harry only wrote 3 and Susan only wrote 4.

For both B and B', *only* in *Susan only wrote 4* is wrongly predicted to be infelicitous since a previous sentence in the sequence, namely *Bill wrote 5,* is not modified by *only,* although given the context ("the rest didn't do so well"), it could be easily felicitously modified by it (as seen by the optionality brackets around *only* there). In other words, given the context, the evaluative inference of *Bill only wrote 5* is true (writing 5 papers is considered less than the norm ('not so well') / less than expected, etc.), but unlike what is predicted, it need not be actually modified by *only* to make the lower *only* felicitous.

More generally, unlike what the explanation and generalization above predict, modification of a sentence by *only* does not seem to necessitate modification of prior sentences by *only*, even if those can be felicitously modified by it. But if this is so, then the infelicity of *only* in e.g. (22a) (as opposed to its felicity of in (23) with 'sequence of *only*') cannot be just attributed to the fact that the previous sentence is not modified by *only*.

---

[10] Notice that what causes infelicity is the presence of a weaker / stronger sentence (respectively) which can appear early in a sequence of sentences, and not directly coordinated with *only/even p*. This is seen in (i):

(i) How many papers did your faculty members write in this period?
    a. *Sean wrote 7, John wrote 2, Bill wrote 5 and Harry (#only) wrote 4*
    b. *Sean wrote 5, John wrote 7, Bill wrote 4 and Harry (#even) wrote 6*

Instead we propose again that the infelicity of e.g. (22a) is because of the presence of an alternative in C which is weaker than the prejacent of *only*, whereas *only* presupposes that *p* is the weakest alrantive in C. This, though, raises again the question of why this infelicity is prevented when the weaker alternative is itself modified by *only* as in (23), and more generally, what precisely happens with sequences of *only* and sequences of *even.*

A preliminary suggestion is made in Greenberg 2019 regarding the 'sequence of *only*' case in (23a). Greenberg hypothesized that "It seems as though the presence of the *only* in the first sentence [of (23a)] 'shields' [the prejacent], and indicates that [it] does not need to be considered a 'contextually relevant' alternative. Thus, the prejacent of the second *only* should not be compared to it, but to another higher alternative" (p. 24).

A similar suggestion can be made for the 'sequence of *even*' case in (24). More generally we hypothesize that once a sentence is a prejacent of an alternative-sensitive scalar operator, it can only be compared to a higher / lower alternative on the relevant scale, and cannot itself play a role as a member of sets of alternatives for other prejacents of the same operator. Thus, such prejacents do not enter into the C sets of other prejacents. In other words, sequences of *only* or of *even*, 'flatten' the scalar relations between their own prejacents, so these are considered as equal in the comparison to alternatives in C.

This kind of hypothesis needs to be motivated and made more precise. If it ends up being on the right track, though, it can help explain felicity contrasts with *even* and *only* of the sort examined above while maintaining a scalar presupposition with universal quantification over alternatives in C for both *only* and *even*. We leave further investigation of this issue to future research.

5.   Deriving mirativity / evaluativity effects of *only*

In section 2 we argued that evaluativity (requiring *p* to be 'small' / less than the standard) should not be hardwired into the entry of *only,* as it is neither a necessary nor a sufficient condition on the felicity of this particle. On the other hand, the tendency of *only* to trigger such evaluative inferences is very prominent. How can such a tendency be explained?

To answer this question we will rely here on the interaction between the scalar presupposition in the **hybrid** entry on *only*, requiring *p* to be the weakest alternative in C and an idea suggested in Krifka's 2000, regarding sentences with *Schön (already,* as in *Lidia is already 3 months old)* and *noch* (*still,* as in *Lidia is still 3 months old*). Krifka discusses the 'mirative' inferences that such particles raise (requiring *p* to be earlier and later than expected, respectively), and like others (e.g. Löbner 1989), argues that they areare cancellable, and hence should not be hardwired. Moreover, Krifka suggests a way to derive such effects. This way relies on the interaction between (a) a semantics with universal quantification, where *p* indicates a point later than all alternatives in C (for *already*) and earlier than it (for *still*), and (b) "a general pragmatic rule, a consequence of the maxim of relevance [namely that] the alternative propositions must be considered reasonable, or entertainable, at the current point in discourse" These, Krifka suggests, lead to the situation where *still* and *already* "express a deviation from expected values in a particular direction" (p. 405).

An explanation along the line of Krifka 2000 can be now developed for the "smallness" effects of *only* (cf. also Orenstein 2016, Liu 2017 for suggestions along this line as well). The idea would be that if indeed all alternatives to the prejacent of *only, p,* in C must be

stronger than it (as in the Guerzoni-style **hybrid** entry in (3c) above), and if indeed all such alternatives must be relevant / entertainable in the context, then the need to accommodate such stronger alternatives in order to construct C leads to considering p less than what is contextually reasonable / entertainable, i.e. as "small".

Moreover, when constructing this set of alternatives, which need to be both stronger than p and contextually entertainable, clashes with existing assumptions about what is reasonable or entertainable in the context, we end up with infelicity. We suggest that this is what happens in e.g. (7c) above (#*John only has 8 kids*), where taking stronger alternatives (e.g. *John has 9 kids, John has 10 kids,* etc..) to be contextually entertainable / relevant clashes with existing common ground assumption regarding the number of kids in typical western contexts.[11]

A novel support for the proposal above is the observation that the cases where evaluativity effects of *only* disappear, like (9) above (i.e. those where *only* is felicitous despite the fact that p is not small, but 'above the standard'), are exactly those where there is discourse salient material (e.g. VPs) which can be used to construct stronger alternatives.

We suggest that this is because, unlike what happens in the null context, here accommodating alternatives into C is not needed, as the speaker can exploit the discourse salient VPs for constructing the alternatives. Crucially, then, in such cases no assumptions about what is or what is not reasonable / expected in the context need to be made, so no 'less than the standard / less than expected' inferences arise.[12]

We conclude, then, that evaluativity of *only* (indicating 'less than the standard') should be indeed not hardwired into its semantics (in contrast to that of *even*, indicating 'more than the standard'). In all its uses *only p* just presupposes that p is weaker than all its focus alternatives in C, which are either accommodated (thus indirectly leading to 'mirativity effects'), or constructed based on salient material (thus allowing a 'mirativity-free' interpretation).

Notice that this conclusion makes *even* and *only* scalar antonyms just in terms of the comparative scalar ordering between p and its alternatives, and not in terms of hardwired evaluativity (i.e. a requirement for a degree which is more than vs. less than what is expected or what the standard is).

6.   Summary

Beaver & Clark 2008 write that:

> In considering the meanings of *only* and *even*, one is tempted to say that they are, in some sense, opposites. Yet is hard to put one's finger on the nature of this intuitive antonymy. ….We suggest that *only* and *even* might best be labeled PRAGMATIC ANTONYMS (p. 71)

---

[11] Iatridou & Zeijlstra 2019 make a similar suggestion regarding the adverbials *in years* and *until-p*. They observe that such adverbials trigger 'beyond expectation inferences' and suggest to derive them from the assumptions (a) that these adverbials denote time spans which are wider than all their and (b) they are contrastively stressed, and hence act as domain wideners, i.e. their 'prejacent' time spans are perceived as wider than what is considered contextually relevant / expected.

[12] See Umbach 2009 for a similar explanation for the absence vs. presence of 'norm-related' effects with comparative *noch*.

One main conclusion of this paper is that adopting what we called a **hybrid** semantics for English *only*, developed in e.g. Guerzoni 2003, allows us to "put our finger" on the nature of the intuitive antonymy referred to in Beaver & Clark, and capture it semantically.[13] Given this suggestion, both particles have scalar presuppositions with universal force, requiring *p* to be weaker (in the case of *only*) and stronger (in the case of *even*) than all other alternatives in C. We argued that adopting this entry is further supported by that fact that it does a better job than the more popular **non-scalar** and **scalar** entries in explaining the infelicity of *only* where salient sentences are weaker than *p* (and the fact that this infelicity is the mirror image of infelicity patterns with *even).*

Importantly, in contrast to this antonymy in scalar ordering (between *p* and its alternatives in C), we argued that *only* and *even* should not be considered 'evalautive' / 'mirative' antonyms. While *even* is indeed a real evaluative particle, presupposing that *p* (as well as its alternatives) indicates a degree which is higher than the salient standard on the relevant scale (Greenberg 2018), the apparent mirror imaged evaluativity effects of *only* (indicating a degree below the standard) is not hardwired, but rather derived from the universal quantification in the **hybrid** and constraints on alternatives in C. This suggestion allowed us to not only to explain the fact that the evaluativity effects of *only* are cancellable, but to also predict in which contexts they can indeed dissappear.

Finally, we showed that the maintaining the universal quantification over alternatives in C in the scalar presupposition of the **hybrid** entry of *only* and in the traditional entry for *even* faces parallel challenges (e.g. with contextually relevant alternatives based on the lexicon, and sequences of *only* and sequences of *even*). We took such challenges to further support the parallel sensitivity of these particles to constraints on constructing the set of alternatives, C, though we emphasized that these constraints should be further studied.

The suggestions above raise many open questions. Here are some of them: (a) What are other challenges for the universal quantification over alternatives in scalar presupposition of *only* and *even*, and what are other potential solutions to such challenges? (b) Can we predict in advance which scalar operator will have hardwired evalautivity (like *even*), and for which operators such effects are only derived (as proposed above for *only*)? (c) What other mechanisms for deriving evalautivity effects are there, besides the one discussed in section 4 above? And which operators use such mechanisms (see e.g. *All* discussed in Homer (to appear), *in years* and *until* discussed in Iatridou & Zeijlstra 2019, for which such effects were reported)? (d) What governs the absence of evaluativity effects with some scalar operators? An example is the absence of such effects with covert *exh* (vs. their presence with the overt *only* (cf. Crnič 2012). Should this be attributed to the fact that, unlike *only*, the C set of *exh* can ignore context and can only use lexicon as a source of constructing alternatives (cf. section 5)? Or perhaps this is due to the fact that the potential alternatives to the prejacnet of *exh* are exhaustified as well, leading to a 'sequence of *exh*' (cf. section 4)? We leave these and other questions to future research.

---

[13] Cf. Zimmermann 2014 for an in-depth discussion of antonymies between *even*-like and *only*-like partices cross-linguisticaly,

References

Alxatib, Sam 2013. 'Only' and Association with Negative Antonyms. Ph.D. thesis, MIT.

Beaver, David, and Brady Clark. 2008. Sense and sensitivity: How focus determines meaning. Oxford: Blackwell.

Chierchia, Gennaro. 2013. Logic in grammar: Polarity, free choice, and intervention. Oxford University Press.

Coppock, Elizabeth and David Beaver. 2014. Principles of the exclusive muddle. Journal of Semantics 31, 371-432.

Crnič, Luka. 2012. Focus particles and embedded exhaustification. Journal of Semantics 30:533-558.

Fox, Danny, and Roni Katzir. 2011. On the characterization of alternatives. Natural Language Semantics 19:87–107.

Greenberg, Yael. 2015. *Even*, comparative likelihood and gradability. In Proceedings of the Amsterdam Colloquium 20, edited by. T. Brochhagen, F. Roelofsen, and N. Theiler, 147–156. Amsterdam: UVA.

Greenberg, Yael. 2016. A novel problem for the likelihood-based semantics of *even*. Semantics and Pragmatics 9: 1–28.

Greenberg, Yael. 2018 A revised, gradability semantics for *even*. Natural Language Semantics 26:51-83

Greenberg, Yael. 2019, Scalarity, Exclusivity, Mirativity / Evaluativity: What (and what doesn't) make 'only' a mirror image of 'even'. Ms. Bar Ilan University

Greenberg, Yael. To appear. An overt *even* operator over covert-based focus alternatives: The case of Hebrew *BIXLAL*. Journal of Semantics.

Guerzoni, Elena. 2003. Why even ask? On the pragmatics of questions and the semantics of answers. PhD dissertation, MIT.

Homer, Vincent, To appear. *That's all.* Proceedings of WCCFL 36.

Horn, Laurence. 1969. A presuppositional analysis of only and even. In The 5th Annual Meeting of the CLS, Chicago Linguistic Society, Chicago. 4:97-108.

Iatridou, Sabine, and Hedde. Zeijstra 2019, The complex beauty of boundary adverbials: *in years* and *until* ms.

Karttunen, Lauri, and Stanley Peters. 1979. Conventional implicature. In Syntax and Semantics 11: Presupposition, ed. D.A. Dinneen and C.-K. Oh, 1–56. New York: Academic Press.

Katzir, Roni, 2014. On the roles of markedness and contradiction in the use of alternatives. in Semantics, Pragmatics, and the Case of Scalar Implicatures. ed. by Salvatore Pistoia Reda (ed.). 40-71. Palgrave

Kay, Paul. 1990. Even. Linguistics and Philosophy 13:59–111.

Klinedinst, Nathan., 2004 Only scalar *only.* A paper presented in Presupposition & Implicature Workshop. Paris.

Klinedinst, Nathan. 2005. Scales and *Only.* Master's thesis, UCLA.

König, Ekkehart. 1991. The Meaning of Focus Particles: A Comparative Perspective. London: Routledge

Krifka, Manfred. 2000. Alternatives for aspectual particles: Semantics of *still* and *already*. In Proceedings of the 26th BLS. ed. by Andrew K. Simpson, 401–412

Liu, Mingming, 2017. Varieties of alternatives: Mandarin focus particles. Linguistics and Philosophy 40: 61-95.

Löbner, Sebastian. 1989. German *schon - erst - noch:* An integrated analysis. Linguistics and Philosophy, 12:167-212.

Orenstein, Dina. 2016. The semantics, pragmatics and focus sensitivity of exclusive particles in Modern Hebrew. PhD dissertation, Bar Ilan University.

Orenstein, Dina. & Yael Greenberg. 2013, A Flexible Exclusive Particle: The Case of the Hebrew *be-sax ha-kol ("all in all")*. In Proceedings of SuB 17. Edited by Emmanuel Chemla, Vincent Homer, Gr´egoire Winterstein. 381-397.

Roberts, Craige. 2011. Only: A case study in projective meaning. In The Baltic International Yearbook of Cognition, Logic and Communication.Volume 6.

Rooth, Mats. 1992. A theory of focus interpretation. Natural Language Semantics 1:75–116.

Umbach, Carla. 2009, Comparative combined with additive particles: The case of German *noch*, proceedings of SuB 13.

Xiang, Yimey, 2019. Function alternations of the Mandarin particle *dou*: Distributor, free choice licensor and 'even', ms. Rutgets.

.Zeevat, Henk. 2008. "Only" as a mirative particle. In Focus at the Syntax-Semantics Interface, ed. by Arndt**.** Riester and Edgar Onea. Working Papers of the SFB 732, Vol. 3. 121-140. University of Stuttgart.

Zimmermann, Malte. 2014. Scalar particles and contrastive topics in English and Vietnamese. Proceedings of IATL 31 (MITWPL 82), edited by Elitzur Bar-Asher Siegal, 123-152. Cambridge, Mass: MIT.

Yael Greenberg
yaelgree@gmail.com