Evaluativity in *even*-comparatives via presupposition accommodation[1]

Ruyue Agnes BI — *Massachusetts Institute of Technology*

Abstract. Sentences involving the scalar particle *even* result in positive or negative inferences when combined with gradable predicates, which is unexpected assuming the classic (un)-likelihood analysis of *even*. I observe that these evaluative inferences arise when (i) the likelihood presupposition of *even* is not already entailed in the common ground, and (ii) the comparison class is not explicitly given. Hence, this paper gives a pragmatic account of the puzzle and demonstrates that the evaluative inferences can be derived systematically with the help of two general pragmatic principles, what I call the "Presupposition Accommodation Condition" and "Alternative-sampling Hypothesis." The basic intuition is that given the pressure to reason in the absence of sufficient evidence – in these cases, without specified reasons for the likelihood orderings – comprehenders make the most out of the information they have, assume normality, and commit to their best guess for the speaker's intended common ground.

Keywords: evaluative inferences, scalar particles, *even*, abductive reasoning, presupposition accommodation, pragmatics

1. Introduction

Discourse participants are constantly negotiating what they take to be in the common ground. The presupposition φ of an utterance S can often be accommodated when the addressee recognizes that the speaker intends φ to be in the common ground. In that case, if φ does not contradict the addressee's beliefs, they adjust their understanding of the common ground accordingly so as to include φ (Stalnaker, 2002). This paper focuses on how the presupposition of the scalar particle *even* is accommodated in comparative sentences, and how abductive reasoning comes into play in deriving their inferences.

When uttered with minimal background, comparatives with *even* give rise to different inferences depending on where main stress falls, as (1) shows:

(1) a. Alex is even taller than [Blake]$_F$. [$\rightsquigarrow$ Both Alex and Blake are *tall*][2]
b. [Alex]$_F$ is even taller than Blake. [$\rightsquigarrow$ Both Alex and Blake are *short*][3]

Implying that the degree to which an individual exhibits a certain property exceeds the contextually determined standard is referred to as an **evaluative inference**. Positive adjectives, such as *tall*, by themselves do not produce evaluative inferences in the comparatives. We cannot infer Alex's or Blake's tallness status from (2) alone:

[1] I am grateful to Christopher Baron, Patrick D. Elliott, Kai von Fintel, Peter Grishin, Martin Hackl, Roumi Pancheva, Nadine Theiler, and anonymous reviewers of WCCFL 39, CLS 57 and SuB 26 for helpful comments. Special thanks to Athulya Aravind, Danny Fox, and Roger Schwarzschild for extensive discussions and detailed feedback at different stages of this project. Parts of this work were presented at the LFRG at MIT and CLS57. All errors are, of course, mine and mine alone.

[2] This positive inference, to the best of my knowledge, is first noted in Greenberg (2015) example (10).

[3] Another way to bring out this intuition is to move *even* to the beginning of the sentence *Even [Alex]$_F$ is taller than Blake* so that we do not rely on *even* to back-associate. Cf. the comparable version of (1a) *Alex is taller than even [Blake]$_F$*. Thanks to Aviv Schoenfeld for this observation at the CLS poster session.

© 2022, Ruyue Agnes Bi. In: Daniel Gutzmann and Sophie Repp (eds.): Proceedings of Sinn und Bedeutung 26, pp. 112–129. University of Cologne.

(2) Alex is taller than Blake.

Hence, the evaluative inference must have been triggered by the addition of *even*. Although the focus sensitivity is a novel observation, positive inferences observed in *even*-sentences with default, sentence-final pitch accent, such as (1a), have been noted in literature (Greenberg, 2015; Daniels and Greenberg, 2020) and taken as one of the main arguments to enrich the meaning of *even* from its classic formulation (Karttunen and Peters, 1979; Rooth, 1992; Chierchia, 2013; a.o.)

(3) **Classic (un-)likelihood analysis of *even***
$[\![even]\!]^{g,c} = \lambda C.\lambda p : \forall q \in C\,[q \neq p \rightarrow p <_{\text{likely}} q].p$
i.e., the prejacent p is the least likely among its alternatives in the context C

to a more involved one by explicitly encoding some version of a positive condition in its semantics. This paper takes a pragmatic approach instead and claims that the classic (un)likelihood meaning of *even* is sufficient, and the evaluative inferences only arise as a default reading in the absence of more specific contexts. The remainder of the paper is organized as follows: in section 2, I give some relevant background on the puzzle, present the details of my proposal, and illustrate the advantages of a pragmatic approach. In section 3, I review the competing semantic analyses on the market and demonstrate why this line of approaches, as they are formulated right now, is not sufficient to capture the full empirical landscape. Section 4 concludes and points out directions for future work.

2. A pragmatic approach

Positive inferences, such as that of (1a), have been noted and discussed quite extensively in literature. A prominent line of research on this topic advocates for an enrichment of the semantics of *even* (Greenberg, 2015, 2018; Daniels and Greenberg, 2020).

In her 2015 paper, Greenberg points out three novel empirical challenges for the classic (un-)likelihood view: (i) examples where *even p* is felicitous despite the prejacent p not being less likely than an alternative q; (ii) examples where p asymmetrically entails its alternatives q, in which case the presupposition of *even* is supposed to be satisfied, yet *even p* is still infelicitous; (iii) unexpected sensitivity to standards of comparison. I focus on the third issue in this paper, although a successful account of *even* should be able to address the first two as well.

Greenberg observes that when *even* associates with comparatives formed with relative adjectives, it gives rise to entailments of their positive forms.

(4) **Greenberg (2015) Observations**
a. (i) The blue tool is [stronger than the red tool]$_F$. But both are weak.
(ii) The blue tool is *even* [stronger than the red tool]$_F$. #But both are weak.
b. (i) Bill is [taller than John]$_F$. But both are short.
(ii) Bill is *even* [taller than John]$_F$. #But both are short.

Greenberg takes these data to suggest that in addition to the comparison between the prejacent p and its alternative q, both p and q also need to have degrees that are at least as high as the **standard** on the contextually determined scale.

I will take a pragmatic approach instead, and argue that supplementing the classic (un-)likelihood analysis of *even* in (3) with two general pragmatic principles can derive these inferences systematically. In a nutshell, I postulate that given the pressure to reason in the absence of sufficient evidence (in these cases, without specified reasons for the likelihood orderings), comprehenders make the most out of the information they have, and commit to their best guess for the speaker's intended common ground.

In what follows, assume the standard terminology in the degree literature, where the subject of a comparative is referred to as the **target**, and the object of *than*-clause in a phrasal comparative as the **standard**. Suppose *even* always takes scope over matrix TP and associates with focus. I adopt the alternative semantics approach (Rooth, 1985) to focus-sensitivity, and assume that a set of alternatives is computed in the context based on where the focus is placed.

2.1. Proposal

I will illustrate the details of my proposal using the minimal pair in (1). According to the classic (un)likelihood analysis of *even* (3), both sentences assert that *Alex is taller than Blake*, but with distinct presuppositions due to the difference in the alternative sets:

(5) Alternatives of (1a) = $\{$Alex is taller than $\pi \mid \pi \in$ ALT(BLAKE)$\}$
Alternatives of (1b) = $\{\pi$ is taller than Blake $\mid \pi \in$ ALT(ALEX)$\}$

The presupposition of *even* mandates that the proposition *Alex is taller than Blake* is the least likely among its alternatives. This is still quite far from the observed evaluative inferences, since it is a statement of likelihood ranking of propositions and does not evoke the "tallness" scale. The common ground can be either of two types, and we will discuss the two cases in turn. The first case is when the common ground already entails this likelihood presupposition. Then no accommodation is needed, and evaluative inferences do not necessarily arise. The prediction is borne out empirically. Consider the scenario in (6) where discourse participants already have clear reasons to believe that *Alex being taller than Blake* is less likely.

(6) *Context: The speaker is talking about the three kids in the Smith family. So far, it has been established that Alex is taller than her twin Aaron and that Blake is their older sibling.*
Alex is even taller than [Blake]$_{\text{F}}$. cf. (1a)

Since common knowledge suggests twins are usually of similar heights and older siblings tend to be taller than the younger ones, the context above already entails that p 'Alex is taller than Blake' is less likely than q 'Alex is taller than Aaron.' The presupposition of *even*, as stated in the classic (un-)likelihood analysis, is satisfied, and, crucially, no accommodation is required. Indeed, these kids can be tall, short, or average height; we do not necessarily infer anything about their height status relative to an external standard from the *even*-sentence.[4]

[4]It is interesting to note that there is a secondary, somewhat "hidden" implicature of (6) – Blake is taller than Aaron. This becomes clearer if we think of the (admittedly, very elaborate) scenario where the speaker is comparing two kids' heights at a time. The kids are asked to stand side-by-side in pairs. The speaker first compared Alex with her twin Aaron and found out that Alex is taller than Aaron, and then compared Alex with her older sibling Blake, and concluded that *Alex is even taller than Blake*. At this point, the relative height between Aaron and Blake is

The second case is when the common ground does not yet entail this likelihood presupposition. The addressee learns that the likelihood rankings are part of the speaker's *intended* common ground and, somewhat unexpectedly, commences abductive reasoning trying to find the most likely justification for the rankings. Discourse participants are not satisfied with simply accommodating the likelihood presupposition on its own; they need to understand more concretely why the speaker would presuppose such differences in likelihood exist in the first place. I claim that in the absence of further context, height contrast is the best and most salient explanation.

Let's adopt a classic formulation of abduction, i.e., the inference to the best explanation:

(7) **Abduction** (Douven, 2017)
Given evidence E and candidate explanations $H_1, ..., H_n$ of E, infer the truth of *that* H_i which best explains E.

For (1a), E is the evidence that *Alex is taller than Blake* is the least likely proposition among its alternatives $\{$Alex is taller than $\pi\}$. There can be many reasons as to why that would be the case. To name a few, let the candidate explanation H_1 be Blake is the oldest in the group of teenagers, H_2 Blake is a mythical giant, H_3 Blake is the only modern human compared to a group of mid-19th century people[5], so on and so forth. Among these explanations, H_i Blake is taller than everyone else, apart from Alex, within the contextually determined group is the one that presumes the least from context, which, I hypothesize, makes it the "best" explanation of E. Hence, via abduction, listeners infer H_i to be true, i.e., Blake is taller than everyone else within the set of the individuals considered in the computation of alternatives. We will refer to this set as the set of **individuals under consideration**, and it forms the comparison class for the interpretation of the relative adjective.

In a broader context, the insight behind this reasoning is what I argue to be a general hypothesis of how presuppositions are accommodated.

(8) **Presupposition Accommodation Condition (PAC)**
In cases where a presupposition is not entailed by the common ground, listeners can accommodate it only if its best explanation is already entailed or can be accommodated simultaneously.

For *even*-comparatives like (1), how the individuals are ordered on the scale of the gradable predicate is assumed by the listener to be *a priori* justification for likelihood contrasts, and thus needs to be accommodated along with the likelihood rankings. The link between degrees and likelihoods is a natural one, considering the *transitive* and *antisymmetric* properties of ordering

still unknown. Finally, the speaker compared these two and discovered that Aaron is actually a little bit taller than Blake as well. Given this context, some native English speakers are asked to evaluate the *even*-sentence uttered at two different time points: (i) before the total ordering is known; (ii) after it has been established in the conversation that Aaron is taller than Blake. It is reported that the sentence is felicitous at the former stage but not at the latter. As argued before, there is no likelihood accommodation needed in this context, but listeners can, nonetheless, draw inferences predicted by the likelihood-to-degree mapping corollary, which I will introduce in (12). A rough derivation is as follows: the alternative set here is $\{$Alex is taller than $\pi \mid \pi \in \{$Aaron, Blake$\}\}$; the relevant $\textsc{g}$ is the property *tall*, $x =$ Blake, $y =$ Aaron, $z =$ Alex. The best explanation for $\mathscr{L}(\text{Alex} >_{\text{TALL}} \text{Blake}) < \mathscr{L}(\text{Alex} >_{\text{TALL}} \text{Aaron})$ is that Blake $>_{\text{TALL}}$ Aaron, and, hence, we assume it to be true. Thanks to Danny Fox (p.c.) for pointing this out.

[5]According to Scientific American, over the last 150 years, the average height of people in industrialized nations has increased approximately 10 centimeters.

relations. Cresswell (1976) gives a formal definition of degrees in set theory as points on a scale – the scale is represented by a relation $>$, which is a set of ordered pairs; and the points on the scale is represented by the field of that relation $\mathfrak{F}(>)$, which is the set of all things that are related in one direction or another to something else. The field of a relation is also often referred to as its dimension (e.g., height, cost, crime rates, temperature).

(9) A DEGREE (of comparison) is a pair $\langle u, > \rangle$, where $>$ is a relation and $u \in \mathfrak{F}(>)$.
(Cresswell, 1976, p. 266)

Cresswell further notes the relation $>$ is usually thought of as at least a partial ordering, from which we can conclude

(10) For a gradable predicate G with responding scale $>_{\text{G}}$, and $x,y,z \in \mathfrak{F}(>_{\text{G}})$,
a. Given $x >_{\text{G}} y$. If $z >_{\text{G}} x$, then $z >_{\text{G}} y$;
b. Given $y >_{\text{G}} x$. If $x >_{\text{G}} z$, then $y >_{\text{G}} z$.

Translating this into **likelihood** $\mathscr{L}$

(11) For a gradable predicate G with responding scale $>_{\text{G}}$, given any $x,y,z \in \mathfrak{F}(>_{\text{G}})$,

$$\begin{cases} x >_{\text{G}} y \rightarrow \mathscr{L}\big(z >_{\text{G}} x\big) \leq \mathscr{L}\big(z >_{\text{G}} y\big) & \text{(a)} \\ y >_{\text{G}} x \rightarrow \mathscr{L}\big(x >_{\text{G}} z\big) \leq \mathscr{L}\big(y >_{\text{G}} z\big) & \text{(b)} \end{cases}$$

Note (11) is a fundamental property of comparatives, independent of any specific behaviors of different types of adjectives.

I suggest that the antecedents serve as the best explanations of their corresponding consequent when there is not already a salient explanation in the context. For example, x being higher on the scale than y, i.e. $x >_{\text{G}} y$, best explains why the likelihood of z being higher on the scale than x is less than or equal to the likelihood of z being higher on the scale than y, $\mathscr{L}\big(z >_{\text{G}} x\big) \leq \mathscr{L}\big(z >_{\text{G}} y\big)$. Hence, given PAC, in order to accommodate the likelihood rankings, listeners need to accommodate the orderings on the relevant scale as well:

(12) **PAC Corollary:** ***Likelihood-to-degree Mapping***
Given a gradable predicate G and its responding scale $>_{\text{G}}$, for $x,y,z \in \mathfrak{F}(>_{\text{G}})$,
a. Accommodating $\mathscr{L}\big(z >_{\text{G}} x\big) < \mathscr{L}\big(z >_{\text{G}} y\big)$ requires that $x >_{\text{G}} y$ is either entailed in the common ground or can be accommodated;
b. Accommodating $\mathscr{L}\big(x >_{\text{G}} z\big) < \mathscr{L}\big(y >_{\text{G}} z\big)$ requires that $y >_{\text{G}} x$ is either entailed in the common ground or can be accommodated.

When listeners hear *even*-comparative sentences uttered out of the blue, they learn that they need to accommodate a certain set of likelihood rankings. With the help of the likelihood-to-degree mapping, they can reach an interim conclusion of where the individual in question sits relative to other members of the comparison class on the relevant scale. In (1a) where the comparison class is not explicitly given and presumably contains more than two individuals, this PAC corollary guarantees Blake is taller than all the other individuals under consideration.

It is important to note that Blake being taller than everyone else within the comparison class does not necessarily mean Blake is tall. Perhaps everyone else in the comparison class is exceptionally short, in which case we still do not know Blake's height status relative to the external standard. To bridge this gap, let us assume that when an *even*-comparative is uttered more or less out of the blue, i.e. when the individuals under consideration are not specified in the discourse, the comparison class is representative of the contextually determined population. This proposal is formalized as follows:

(13) **Alternative-sampling Hypothesis (ASH)**
When the alternative set is not explicitly specified, assume that individuals included in the computation of alternatives form a representative sample of the contextually determined relevant population with respect to G.

Essentially, all (13) claims is that interlocutors always assume normality unless told otherwise.

Now we have all the necessary pieces to give a complete derivation of the positive inference in (1a) when uttered in a context that does not already support the presupposition of *even*. The classic (un-)likelihood analysis of *even* presupposes that *Alex is taller than Blake* is the least likely among its alternatives $\{$Alex is taller than $\pi \mid \pi \in \mathrm{ALT}(\textsc{Blake})\}$. First, the likelihood-to-degree mapping guarantees that *Blake is taller than everyone else in the comparison class*. Second, since the height of a population is normally distributed and, by Alternative-sampling Hypothesis, the comparison class reflects a representative sample, for Blake to be at least as tall as everyone else in that set, Blake is taller than the population median. Hence, we get the inference that Blake is *tall*. Third, since *even* asserts that its prejacent is true, Alex is taller than Blake, which, in turn, indicates that Alex is tall as well. Notice the inference that *Blake is tall* is of a different status as that of *Alex is tall*, with the former derived entirely from presupposition and the latter additionally relying on the prejacent being TRUE.

The reasoning process is sketched in the figure below – the shaded box is what we have from the semantics of *even* alone, and I propose that listeners need to go through three additional steps to arrive at the evaluative inferences. Justification for each step is noted above the arrow.

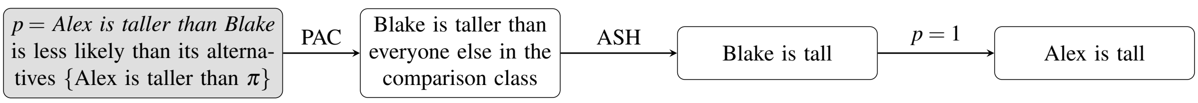


Figure 1: Pragmatic reasoning process for (1a)

The reversed, negative inference of (1b) falls out straightforwardly from the same reasoning process. Let us assume that negative gradable predicates share the same fields as their positive counterparts but with opposite relations. Starting with the alternative set $\{\pi$ is taller than Blake $\mid \pi \in \mathrm{ALT}(\textsc{Alex})\}$, the classic (un-)likelihood analysis requires that p 'Alex is taller than Blake' is the least likely among its alternatives. The relevant G is again the property *tall*, $x =$ Alex, $y =$ everyone else considered in the alternatives, and $z =$ Blake. By part (b) of the likelihood-to-degree mapping, to accommodate the said likelihood rankings, we accommodate at the same time that Alex is the shortest within the comparison class. By the Alternative-sampling Hypothesis, since the comparison class is a representative sample of a normal distribution, Alex is shorter than the population median. Lastly, assuming the prejacent to be TRUE, Alex, albeit short, is taller than Blake, so Blake must be short as well.

Notice an essential difference between the two derivations: for (1a), comprehenders come to a conclusion about Blake's height status first, while for (1b), they do so about Alex's height status

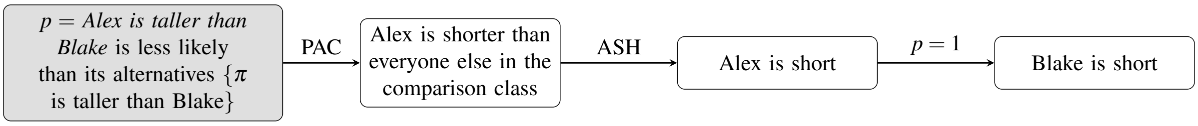


Figure 2: Pragmatic reasoning process for (1b)

first. Therefore, this analysis predicts that when we have reasons to reject the prejacent or when the prejacent is not automatically assumed to be true, for example, in (1)'s corresponding polar questions, we infer the height status of distinct individuals. This seems to be consistent with native speakers' judgment:

(14) a. Is Alex even taller than [Blake]$_F$?
b. If Alex is even taller than [Blake]$_F$, Alex will be able to reach the top shelf.

From (14), listeners infer that Blake is tall, but not necessarily that Alex is tall. The sentences are felicitous even when Alex's height is unknown to the speaker.

(15) a. Is [Alex]$_F$ even taller than Blake?
b. If [Alex]$_F$ is even taller than Blake, Blake cannot go on the roller coaster ride.

Similarly, from (15), listeners infer that Alex is short, but not necessarily that Blake is short.

2.2. Adjective type is irrelevant

As noted before, the property that the likelihood-to-degree mapping relies on is fundamental to all partial ordering relations, and thus to all types of gradable predicates. The minimal pair in (1) features a quintessential positive adjective, *tall*. For completeness, I will walk through the derivations for a comparable set of examples containing the negative adjective *cheap* and show that the exact same reasoning applies.

Consider the minimal pair in (16) uttered out of blue:

(16) a. Pomelos are even cheaper than [pomegranates]$_F$. [$\rightsquigarrow$ Both fruits are *cheap*]
b. [Pomelos]$_F$ are even cheaper than pomegranates. [$\rightsquigarrow$ Both fruits are *expensive*]

For (16a), the presupposition of *even* requires that the proposition *pomelos are cheaper than pomegranates* is the least likely among its alternatives {pomelos are cheaper than $\varphi \mid \varphi \in$ ALT(POMEGRANATES)}. The relevant G is the property *cheap*, $x =$ pomegranates, $y = \varphi$, $z =$ pomelos. By (a) of the likelihood-to-degree mapping corollary, to accommodate $\mathscr{L}(z >_{\text{CHEAP}} x) < \mathscr{L}(z >_{\text{CHEAP}} y)$, we need to accommodate $x >_{\text{CHEAP}} y$, i.e., pomegranates are cheaper than everything else under consideration. It follows that pomegranates are the cheapest within comparison class. By the Alternative-sampling Hypothesis, since the comparison class forms a representative sample, pomegranates having a price lower than sample median indicates it is also below population median. Pomegranates are cheap. Since the sentence asserts that the price of pomelos is lower than that of pomegranates, pomelos are cheap as well.

Similarly, for (16b), p 'pomelos are cheaper than pomegranates' is the least likely among its alternatives {φ are cheaper than pomegranates $\mid \varphi \in$ ALT(POMELOS)}. Applying (b) of the likelihood-to-degree mapping, where the relevant G is the property *cheap*, $x =$ pomelos, $y = \varphi$, and $z =$ pomegranates, we deduce that in order to accommodate $\mathscr{L}(x >_{\text{CHEAP}} z) < \mathscr{L}(y >_{\text{CHEAP}} z)$, accommodate at the same time that $y >_{\text{CHEAP}} x$, i.e., everything else under consideration is cheaper than pomelos. This means pomelos have a price higher than that of everything else in the comparison class. By the Alternative-sampling Hypothesis, pomelos have a price higher than not only the sample median but also the population median; pomelos are expensive. The sentence asserts that pomelos are cheaper than pomegranates. Therefore, pomegranates must be expensive too.

We arrive at the correct conversational implicatures without needing to stipulate anything special for negative adjectives. The arguments go through as expected.

Furthermore, we observe that absolute adjectives, such as *safe* and *dirty*, show the exact same inference patterns in *even*-comparatives, despite differing from relative adjectives, such as *tall* and *cheap*, in numerous ways (Rotstein and Winter 2004; Kennedy and McNally 2005; Kennedy 2007; a.o.).

(17) a. District A is even safer than [District B]$_{\text{F}}$. [$\rightsquigarrow$ Both districts are *safe*]
b. [District A]$_{\text{F}}$ is even safer than District B. [$\rightsquigarrow$ Both districts are *dangerous*]

(18) a. The study is even dirtier than [the kitchen]$_{\text{F}}$. [$\rightsquigarrow$ Both rooms are *dirty*]
b. [The study]$_{\text{F}}$ is even dirtier than the kitchen. [$\rightsquigarrow$ Both rooms are *clean*]

Absolute adjectives, unlike relative adjectives, do not exhibit the same level of context sensitivity. What is considered *safe* or *dirty* varies little from context to context. While an individual can be *tall* for regular people but *short* for basketball players, a district, if it is almost free of crimes, will count as *safe* regardless of the criteria and can never be called *dangerous*. As a result, absolute and relative adjectives show distinct patterns when combined with degree modifiers (19) or with an overt *for*-phase specifying the comparison class (20), and with respect to entailments (21).

(19) Compatibility with proportional modifiers
a. District A is *mostly* safe.
b. ??Alex is *mostly* tall.

(20) Compatibility with *for*-phrase
a. ?District A is safe for a financial district[6].
b. Pomelos are cheap for imported fruits.

(21) Entailment patterns
a. District A is not safe. $\models$ District A is dangerous.
Alex is not tall. $\not\models$ Alex is short.
b. The study is dirtier than the kitchen. $\models$ The study is dirty.
Pomelos are cheaper than pomegranates. $\not\models$ Pomelos are cheap.

Kennedy (2007) attributes the absolute/relative distinction to whether the scale used has a maximal or minimal endpoint. Lassiter and Goodman (2013) further generalizes this boundedness property and suggests that "prototypical relative interpretations arise with priors with a relatively mild rate of change and little or no mass on the endpoints, while prototypical absolute

[6]This sentence becomes felicitous under the assumption that financial districts are generally dangerous.

interpretations arise with priors in which a significant portion of the prior mass falls close to an upper or lower bound" (p. 599). Regardless of the specific analysis, the general consensus is that the scales for relative adjectives and the ones for absolute adjectives show disparate characteristics. Since my proposal replies only on the minimal assumption of the scale being a partial ordering, it is appealing for the sake of theoretical parsimony.

3. Previous analyses

In this section, I will review the existing semantic approaches proposed in Greenberg (2015, 2018) and Daniels and Greenberg (2020), and present arguments to why a pragmatic account is more desirable in accounting for these sets of data.

Generally speaking, their work advocates for an enrichment of the semantics of *even* but branches into two distinctive analyses, varying whether the scale of likelihood is retained in the revised meaning of *even*.

3.1. Gradability-based account

Greenberg (2015, 2018) approach, building on an intuition in Rullmann (2007), proposes two major revisions to the classic (un-)likelihood analysis (3): first, instead of imposing restrictions on likelihood, the presupposition of *even* is stated on a contextually determined scale G; second, the standard of G is introduced as part of the presupposition so that the observed evaluative inferences can be straightforwardly captured:

(22) **Gradability-based analysis of *even*** (Greenberg, 2015, 2018)

$$\begin{aligned}
&[\![even]\!]^{g,c} = \lambda C.\lambda p.\lambda w : \forall q \in C\big[q \neq p \rightarrow \\
&\quad \forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_2 \in p \wedge w_1 \in [q \wedge \neg p]] \rightarrow \\
&\quad [max(\lambda d_2.G(d_2)(x)(w_2))) > max(\lambda d_1.G(d_1)(x)(w_1)) \quad \text{Superlative Condition (a)} \\
&\quad \wedge \\
&\quad max(\lambda d_1.G(d_1)(x)(w_1)) \geqslant \mathbf{Stand}_G\,]\big]. \quad \text{Positive Condition (b)} \\
&\quad p(w) = 1 \quad \text{Assertion}
\end{aligned}$$

where x is some salient individual denoted by a non-focused or a contrastive topic constituent in the clause p, C is the set of alternatives, and G is a contextually supplied gradable property.

In prose, *even* presupposes that with respect to an individual x and a gradable property G,the following two conditions hold: (a) x's maximal degree on the scale associated with G is higher in all accessible p worlds than in all accessible q-and-not-p worlds, and (b) in all accessible q-and-not-p worlds, x's degree on G is at least as high as the standard of G.

The revised meaning (22) can easily explain the contrasts between (23a) and (23b), which notably poses a problem for the classic analysis[7]:

(23) *Context: John is an accountant, working in a formal office environment.*

[7] I do not have a satisfactory response to (23) at the moment. These cases pose a very interesting challenge to a pragmatic approach, and I hope to be able to address them in future work.

a. John wore a colorful T-shirt to work yesterday, and he even wore [a funny old hat]$_F$.
b. #John wore his usual white shirt to work yesterday, and he even wore [a funny old hat]$_F$.

Presumably, John wearing a funny old hat is less likely than him wearing his usual shirt, which should have been sufficient to satisfy the presupposition of *even* in the classic analysis, and (23b) is thus predicted to be felicitous, contrary to speakers' intuition. However, if we adopt the proposal in (22), assuming the relevant gradable property G is being unexpected, the alternative $q =$ *John wore his usual white shirt* does not surpass the threshold of what is considered unexpected. This violates the Positive Condition and leads to a presupposition failure.

How this analysis deals with the puzzle we see in (1) is much less clear. Consider (1a) first, repeated below:

(1a) Alex is even taller than [Blake]$_F$. [$\rightsquigarrow$ Both Alex and Blake are **tall**]

Following Greenberg (2018), assume *even* associates with the Degree Phrase *-er than Blake*[8], and the set of alternatives C is as in (24):

(24) $C_{(1a)}$ = { Alex is as tall as Blake, Alex is taller than Blake }
$= \{ \max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) \geq \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake})),$
$\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) > \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake})) \}$

Suppose the gradable predicate G measures degrees to which x is tall. Note that, as Greenberg (2018) mentions in footnote 19, the value of G does not have to be the gradable property the comparative is based on. This analysis leaves some flexibility to the interpretation of G, but also runs the risk of overgenerating. The scalar presupposition of the sentence (1a) is then

(25) Presupposition of (1a) per Greenberg (2018)
$\forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_2 \in [\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) > \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake}))]$
$\wedge w_1 \in [\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) = \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake}))]] \rightarrow$
$[max(\lambda d_3.\text{TALL}(d_3)(\text{Alex})(w_2))) > max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})(w_1))$ Superlative Cond
$\wedge max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})(w_1)) \geqslant \text{Stand}_{\text{TALL}}]]$ Positive Cond

Only when Blake's height is known or fixed in the context, is the superlative condition guaranteed to be trivially satisfied – Alex's degree of tallness in all accessible worlds where she is taller than Blake (p worlds) is higher than in all worlds where she is precisely Blake's height (*q-and-not-p* worlds). The positive condition further requires Alex's degree of tallness in the worlds where she is the exact same height as Blake to be at least as high as the standard for tallness. It follows that Blake is tall, and since Alex is taller than Blake, Alex is qualified as tall as well.

The reasoning goes through roughly as expected, but the additional assumption of the height of Blake being fixed seems too restrictive. However, more pressing issues arise when we consider the case with the main focus on the subject

(1b) [Alex]$_F$ is even taller than Blake. [$\rightsquigarrow$ Both Alex and Blake are *short*]

[8] I assume a phrasal comparative structure here for the ease of illustration. This is not crucial for Greenberg's analysis. The analysis is compatible with a clausal comparative construction.

Assume *even* associates with the subject Alex, and the set of alternatives C is then

(26) $C_{(1b)} = \{$ ALT(ALEX) is taller than Blake, Alex is taller than Blake $\}$

For the purpose of demonstration, let us assume $\{x|x \in \text{ALT}(\text{ALEX})\}$ in this context is a singleton set containing only the individual Dominique. Suppose the gradable predicate G again measures degrees of tallness. The scalar presupposition of the sentence (1b) is thus

(27) Presupposition of (1b) per Greenberg (2018) (attempt 1)

$$\forall w_1, w_2[w_1Rw \wedge w_2Rw \wedge w_2 \in [\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) > \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake}))]$$
$$\wedge w_1 \in [\max(\lambda d_3.\text{TALL}(d_3)(\text{Dominique})) > \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake})) \wedge$$
$$\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) \leq \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake}))]] \rightarrow$$
$$[max(\lambda d_2.\text{TALL}(d_2)(\text{Blake})(w_2))) > max(\lambda d_1.\text{TALL}(d_1)(\text{Blake})(w_1))$$
$$\wedge max(\lambda d_1.\text{TALL}(d_1)(\text{Blake})(w_1)) \geqslant \text{Stand}_{\text{TALL}}]]$$

In other words, the first conjunct requires that Blake's degree of tallness in all accessible worlds where she is shorter than Alex (*p* worlds) is higher than in all worlds where she is shorter than Dominique but taller than Alex (*q-and-not-p* worlds). Assuming Alex's height is known and fixed, it is always false, which wrongly predicates the sentence to be infelicitous. In an attempt to resolve this issue, I will try reversing the scale, accommodating a G$'$ measuring degrees of *shortness*. Note that it is unclear why this rescue strategy can be employed, and whether it is always systematically available.

(28) Presupposition of (1b) per Greenberg (2018) (attempt 2)

$$\forall w_1, w_2[w_1Rw \wedge w_2Rw \wedge w_2 \in [\max(\lambda d_1.\text{SHORT}(d_1)(\text{Alex})) > \max(\lambda d_2.\text{SHORT}(d_2)(\text{Blake}))]$$
$$\wedge w_1 \in [\max(\lambda d_3.\text{SHORT}(d_3)(\text{Dominique})) > \max(\lambda d_2.\text{SHORT}(d_2)(\text{Blake})) \wedge$$
$$\max(\lambda d_1.\text{SHORT}(d_1)(\text{Alex})) \leq \max(\lambda d_2.\text{SHORT}(d_2)(\text{Blake}))]] \rightarrow$$
$$[max(\lambda d_2.\text{SHORT}(d_2)(\text{Blake})(w_2))) > max(\lambda d_1.\text{SHORT}(d_1)(\text{Blake})(w_1))$$
$$\wedge max(\lambda d_1.\text{SHORT}(d_1)(\text{Blake})(w_1)) \geqslant \text{Stand}_{\text{SHORT}}]]$$

The first conjunct now requires that Blake's degree of *shortness* in all accessible worlds where she is shorter than Alex (*p* worlds) is higher than in all worlds where she is shorter than Dominique but taller than Alex (*q-and-not-p* worlds), which, assuming Alex's height is constant across all possible worlds, is trivially met. The second conjunct requires Blake's degree of shortness in the worlds where Alex is taller than her to surpass the standard for shortness. Crucially, the presupposition in (28) does not impose any condition on Alex's height status. We can only infer that Blake is short, but not that Alex is also short.

The gradability-based account of *even* not only is unsuccessful in capturing the complete inferences of comparative sentences with subject focus, but also imposes arbitrary restrictions on what needs to be known in the conversation background. The precise height of the focused element, Blake in (1a) and Alex in (1b), needs to be known for the derivation to go through. In addition, since this approach leaves open how the relevant G is determined, it remains a mystery why the scale is reversed for the minimal pair (1).

3.2. Adding positive condition to the classic analysis

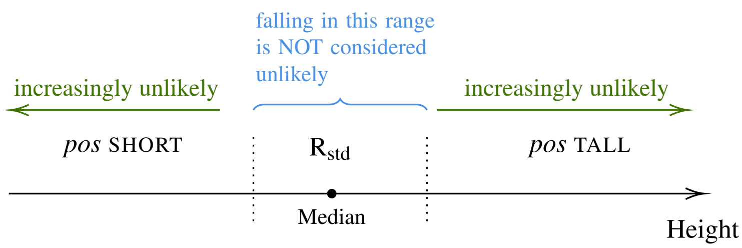
In the same spirit and as an extension of the analysis discussed in the last subsection, Daniels and Greenberg (2020) talk proposes to address the standard sensitivity of *even* in comparative sentences with object focus by hardwiring a positive condition to its classic (un-)likelihood meaning.

(29) **Revised Comparative (un)Likelihood** (Daniels and Greenberg, 2020)
A proposition *even p* is felicitous only if the following conditions hold:
a. **Superlative Condition**: p is $<_{\text{likely}}$ q , for all q in C, where C is the set of contextually restricted focus alternatives;
b. **Positive Condition**: p and q are $>$ R_{Std} for (un)likelihood, where R_{Std} is the standard range, i.e., both p and its alternatives q are unlikely.

Since the original wording in the handout left it unclear, let us assume the weakest, existential version of the Positive Condition. In other words, there exists a salient alternative q such that both the prejacent p and q are unlikely.

As the authors argue, in order to translate (un)likelihood to other scales, we need to crucially assume standards that are computed based on the distributions of degrees in the comparative class. By default, the distributional standard is located at or around the median value (Solt (2011)), and the degrees to which the individuals in the comparison class possess some property cluster around the standard. Here are a summary of the assumptions adopted and a figure demonstrating these assumptions on the scale of height.

(30) **Daniels and Greenberg's (2020) assumptions about scales of gradable predicates**
a. the standard range R_{Std} is an interval on the scale;
b. antonyms sit on the same scale, occupy the opposite sides outside of R_{Std};
c. having degrees within R_{Std} $>_{\text{likely}}$ having degrees outside R_{Std};
d. having degrees within R_{Std} cannot be unlikely;
e. variations within R_{Std} are allowed, i.e., it is possible to have a total ordering of the individuals fall within R_{Std}, but these differences are negligible, and thus not considered unlikely.



For (1a), the alternative set $C = \{$Alex is taller than Blake, Alex is taller than Cassie ...$\}$. The authors assume that the Superlative Condition inherited from the classic comparative (un)likelihood analysis guarantees the ordering Alex $>_{\text{tall}}$ Blake $>_{\text{tall}}$ Cassie. The additional Positive Condition requires both $p =$ *Alex is taller than Blake* and $q =$ *Alex is taller than Cassie* be unlikely. Since an individual can be tall, short, or neither, we have $3 \times 3 = 9$ possible cases regarding Alex's and Blake's height status, assuming they are in the same comparison class. These are examined separately and the eliminated ones crossed out.

Alex	Blake	
tall	*tall*	?
tall	*average*	?
~~*tall*~~	~~*short*~~	since it's not unlikely to be taller than a short individual, *p* is not unlikely
~~*average*~~	~~*tall*~~	contradicts Alex $>_{\text{tall}}$ Blake
~~*average*~~	~~*average*~~	following assumption (30e), *p* is not unlikely
~~*average*~~	~~*short*~~	since it's not unlikely to be taller than a short individual, *p* is not unlikely
~~*short*~~	~~*tall*~~	contradicts Alex $>_{\text{tall}}$ Blake
~~*short*~~	~~*average*~~	contradicts Alex $>_{\text{tall}}$ Blake
~~*short*~~	~~*short*~~	since it's not unlikely to be taller than a short individual, *p* is not unlikely

The authors acknowledge that the remaining two cases need further examination: (i) If Alex is tall and Blake is neither tall nor short, one could argue *p* is not unlikely if we assume that, within a default distribution, the existence of individuals having degrees outside the standard range is expected. This is another, albeit sensible, assumption that needs independent support. (ii) If both Alex and Blake are tall, since variations are expected among the tall individuals, *p* is not necessarily unlikely. Their proposal would then wrongly predicate a presupposition failure.

Daniels and Greenberg's proposal is a reasonable and straightforward attempt to address the puzzle of standard sensitivity in *even*-comparatives, but it appears to suffer from several issues.

First, the analysis takes for granted that the likelihood ranking of the propositions *p* and *q* directly translates to where the individuals sit relative to one another on the relevant scale. In fact, nothing in our current system guarantees this mapping. In theory, there can be many reasons as to why *p* is less likely than *q*. Consider (31) in the given context:

(31) *Context: Stephanie was in a 10k running race yesterday. She did better than a professional runner from Boston, and*
She even did better than a former marathon winner.

Stephanie running faster than a marathon winner is less likely than her running faster than just any professional runner. However, it does not necessarily mean that the professional runner performed worse than the former marathon winner in this particular race. This example reiterates a similar point made by the sibling scenario in (6). A separate, well-defined machinery is needed to explain why and how likelihoods of propositions are translated to degrees on a scale.

Second, the derivation relies on the assumptions in (30), which are satisfied by relative adjectives but not absolute adjectives (Kennedy, 2007; Toledo and Sassoon, 2011; Lassiter and Goodman, 2013). Hence, the same reasoning does not apply to comparable sentences containing absolute adjectives. Daniels and Greenberg would have to explain the similar inference patterns observed with absolute adjectives (see (17) and (18)) by either proposing a separate mechanism or revising the set of assumptions.

In summary, these semantic approaches fail to capture the complete inference patterns observed in *even*-comparatives, and are unsatisfactory conceptually considering the stipulative assumptions. More importantly, they cannot easily explain the systematic scale reversal effect triggered by the change in focus placement.

Contingent on the mechanism assumed for setting the standard of a gradable predicate, semantic approaches seem to predict that, since the positive condition is hardwired, evaluative inferences always arise in comparative sentences with *even*. This is not the correct prediction, as we have seen in our discussion of (6).

Moreover, similar evaluative inferences are seen in other scalar particles, such as *at least* and *still*, and it would be ideal to have a uniform analysis instead of enriching their semantics individually. I will return to this in the next section.

(32) a. The study is at least cleaner than [the kitchen]$_F$. [$\rightsquigarrow$ Both rooms are *dirty*]
b. Simon is taller still than Paige. [$\rightsquigarrow$ Both Simon and Paige are *tall*]

4. Conclusion and discussion

In conclusion, even though deriving evaluative inferences in *even*-comparatives seems straightforward at first and is sometimes taken for granted, there are, in fact, general pragmatic principles at work, and it showcases how discourse participants make the best use of resources provided by context when certain aspects of the information state are left unspecified.

More specifically, I have proposed two general cooperative principles governing how rational agents reason about language, the *Presupposition Accommodation Condition* and the *Alternative-sampling Hypothesis*. Together, they explain the evaluative inferences that arise in comparatives with *even*. This analysis accounts for the optionality of the evaluative inferences in these sentences and straightforwardly derives the puzzling scale reversal with minimal additional assumptions.

4.1. Motivation for the principles

The core insight explored in this paper concerns abductive reasoning, which has been widely studied in AI (Hobbs et al., 1993), philosophy (see Douven, 2017 for a brief survey) and psychology (Harman, 1965; Lombrozo, 2016; Wilkenfeld and Lombrozo, 2015). A growing body of work in experimental psychology reveals that people have a strong, systematic inclination to favor explanations that are *simple*, *broad*, and *consistent with their prior beliefs*. More recently, abduction has also been utilized to account for various cancellable inferences in formal semantics and pragmatics. Gyarmathy and Altshuler (2020) proposes an abductive framework to derive (non-)culminating accomplishment inferences observed in the Hindi perfective and the Russian imperfective. Gyarmathy (2015) applies abduction to explain an existential presupposition of certain achievement predicates, while Piñón (2011) uses it to arrive at the *actuality entailments* of ability modals.

Contextualized in the broader literature, the Presupposition Accommodation Condition, with abductive reasoning at its core, can be considered one mechanism giving rise to the *reasons, causes, consequences, and concurrences* category of the bridging inferences (Clark, 1975)[9]. Clark defines *bridging* in terms of when a listener cannot find an antecedent "directly in memory," they construct the intended antecedent from what they already know, with the assumption

[9]Thanks to SuB reviewer #1 for the suggestion.

that the speaker abides by the Given-New Contract.

(33) **Given-New Contract** (Clark, 1975, p.170)
The speaker agrees to try to construct the Given and New information of each utterance in context
a. so that the listener is able to compute from memory the unique Antecedent that was intended for the Given information, and
b. so that he will not already have the New information attached to the Antecedent.

Clark goes on to suggest that the bridge built should be the shortest possible that is consistent with the Given-New Contract. In other words, "the listener takes as the intended implicature the one that requires the fewest assumptions, yet whose assumptions are all plausible given the listener's knowledge of the speaker, the situation, and facts about the world" (p. 173). In the case of *even*-comparatives, bridging means finding the simplest explanation for the likelihood presupposition. The listener uses the background knowledge about the key properties of partial orderings to bridge from the presupposed likelihood ranking to what they take to be the intended antecedent – differences on the relevant scale.

With respect to the Alternative-sampling Hypothesis, there have been many proposals in the *focus* literature that assume focus-sensitive particles operate on a restricted set of alternatives[10], i.e., a subset of the logical space that is considered "reasonable, or entertainable, at the current point in the discourse" (Krifka, 2000, p. 405).

Krifka (2000) proposes that the presupposition of an aspectual particle, such as *already* or *still*, imposes a restriction on the set of alternatives to the prejacent – specifically, the alternatives must be ordered and the prejacent must be a maximal or minimal element. Roughly speaking, *already* presupposes that the "valid" alternatives are ranked lower than the prejacent on the relevant ordering, while *still* presupposes that the valid alternatives are ranked higher.

Beaver and Clark (2008) develop a similar idea when accounting for the mirative function of *only*. They argue that hearer's expectation of something stronger than the prejacent being true is an essential part of the meaning of *only*, and capture this by presupposing the sole alternatives "left open" are the ones that are at least as strong as the prejacent. As a result, accommodation, if necessary, involves removing invalid alternatives.

Both Krifka (2000) and Beaver and Clark (2008) assume a general concept of *alternatives under consideration* when analyzing data from two different empirical domains. The Alternative-sampling Hypothesis proposed in this paper is along the same lines.

4.2. Extending to *at least*-comparatives

Returning to the observation that similar evaluative inferences can be observed for *at least* in comparative sentences, I will briefly discuss how my proposal might shed some light on the analyses of what Nakanishi and Rullmann (2009) refer to as "concessive *at least*".

It has been noted that *even* under negation invokes a scale reversal in inferences (Karttunen and Peters, 1979). This is indeed what we observe in comparatives below:

[10] I would like to thank SuB reviewer #1 for pointing me to this body of work.

(34) a. Kimberly is even taller than [Simona]$_F$. [$\rightsquigarrow$ Both are *tall*]
b. Kimberly is not even taller than [Simona]$_F$. [$\rightsquigarrow$ Both are *short*]

(35) a. [Kimberly]$_F$ is even taller than Simona. [$\rightsquigarrow$ Both are *short*]
b. [Kimberly]$_F$ is not even taller than Simona. [$\rightsquigarrow$ Both are *tall*]

Under the NPI theory of *even* (Rooth, 1985; Rullmann, 1997; Giannakidou, 2007)[11], the scalar particle *even* is lexically ambiguous between an "ordinary" meaning and a NPI meaning licensed in the scope of a downward-entailing operator. The classic (un-)likelihood analysis (3) states that ordinary *even* presupposes its prejacent is the least likely among its alternatives, and since NPI *even* is associated with the opposite side of the scale to ordinary *even*, we have

(36) $[\![even_{\text{NPI}}]\!]^{g;c} = \lambda C.\lambda p : \forall q \in C\,[q \neq p \rightarrow q <_{\text{likely}} p].p$
i.e., the prejacent p is the **most** likely among its alternatives in the context C

The evaluative inferences in (34) and (35) can be straightforwardly derived following the analysis proposed in section 2, assuming the LF in (37).

(37) [$_{\text{TP}_2}$ *not* [*even*$_{\text{NPI}}$ [$_{\text{TP}_1}$ *Kimberly is taller than Simona*]]]

For (34), the meaning of NPI *even* in (36) presupposes that TP$_1$ *Kimberly is taller than Simona* is more likely than its alternatives $\{$Kimberly is taller than $\pi \mid \pi \in$ ALT(SIMONA)$\}$. Because negation is a **presupposition hole** which allows the presuppositions of its complement to pass through unchanged (Horn, 1969), TP$_2$ carries the same presupposition. Note the relevant G is the property *tall*, $x =$ everyone else considered in the alternatives, $y =$ Simona, and $z =$ Kimberly. By part (a) of the likelihood-to-degree mapping, to accommodate the said likelihood rankings, we accommodate at the same time that everyone else in the set of individuals under consideration is taller than Simona, i.e., Simona is shorter than everyone else in the comparison class. By the Alternative-sampling Hypothesis, since the comparison class is a representative sample of a normal distribution, Simona is shorter than population median. Finally, assuming the proposition *Kimberly is not even taller than Simona* is TRUE, Kimberly is not taller than Simona, so Kimberly must be short as well. The exact same reasoning applies to (35) except now the alternative set triggered is $\{\pi$ is taller than Simona $\mid \pi \in$ ALT(KIMBERLY)$\}$.

Furthermore, Rullmann and Nakanishi (2009) notes a connection between NPI *even* and *at least*. First, NPI *even* and *at least* can often be used almost interchangeably:

(38) If you answer *at least* / *even* one question correctly, you'll pass.
(Rullmann and Nakanishi, 2009, p. 12)

Second, some crosslinguistic evidence appears to suggest that *at least* and *even* are related:

(39) Rullmann and Nakanishi (2009), p. 17
a. Dutch *zelfs maar* (even only) = NPI 'even'
b. Japanese *-dake -demo* (only even) = 'at least'

[11]NPIs (negative polarity items) are lexical items that are only licensed in negative, downward entailing, or non-veridical environments. I believe we will arrive at the same result following scope theory of *even*, and hence, at least for the time being, I will remain neutral in this debate.

Although there are subtle differences in meaning between NPI *even* sentences and their *at least* counterparts, a natural question to ask is whether similar inference patterns are observed in comparative sentences with *at least*.

(40) a. Kimberly is at least taller than [Simona]$_F$. [$\rightsquigarrow$ Both are *short*]
b. At least [Kimberly]$_F$ is taller than Simona. [?$\rightsquigarrow$ Both are *tall*]

Unfortunately, the intuition is not very clear here. The epistemic reading of *at least*, which signals the speaker is uncertain about the exact height of Kimberly in (40a) or the exact height of Simona in (40b), is so conspicuous that it is difficult to access the intended concessive reading without a specific context. This significant confound aside, a few native speakers I consulted did report the judgment above. It would be interesting to investigate why the judgment for (40a) is considerably clearer than that for (40b), and whether the pragmatic analysis proposed in this paper can be extended to account for these data.

References

Beaver, D. I. and B. Z. Clark (2008). *Sense and Sensitivity: How Focus Determines Meaning*. John Wiley & Sons.

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

Clark, H. H. (1975). Bridging. In B. Nash-Webber and R. Schank (Eds.), *Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing*, TINLAP '75, USA, pp. 169–174. Association for Computational Linguistics.

Cresswell, M. J. (1976). The Semantics of Degree. In B. H. Partee (Ed.), *Montague Grammar*, pp. 261–292. Academic Press.

Daniels, M. and Y. Greenberg (2020). Even and Standards of Comparison. *Sinn und Bedeutung 25* talk handout.

Douven, I. (2017). Abduction. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017 ed.). Metaphysics Research Lab, Stanford University.

Giannakidou, A. (2007). The landscape of EVEN. *Natural Language & Linguistic Theory 25*(1), 39–81.

Greenberg, Y. (2015). *Even*, comparative likelihood and gradability. In T. Brochhagen, F. Roelofsen, and N. Theiler (Eds.), *Proceedings of the 20th Amsterdam colloquium*, pp. 147–156.

Greenberg, Y. (2018). A revised, gradability-based semantics for even. *Natural Language Semantics 26*(1), 51–83.

Gyarmathy, Z. (2015). Culminations and presuppositions. In T. Brochhagen, F. Roelofsen, and N. Theiler (Eds.), *Proceedings of the 20th Amsterdam colloquium*, pp. 167–176. ILLC, Department of Philosophy, University of Amsterdam Amsterdam.

Gyarmathy, Z. and D. Altshuler (2020). (Non)culmination by abduction. *Linguistics 58*(5), 1373–1411.

Harman, G. H. (1965). The Inference to the Best Explanation. *The Philosophical Review 74*(1), 88–95.

Hobbs, J. R., M. E. Stickel, D. E. Appelt, and P. Martin (1993). Interpretation as abduction.

Artificial Intelligence 63(1), 69–142.
Horn, L. R. (1969). A presuppositional analysis of only and even. In R. I. Binnick, A. Davison, G. M. Green, and J. L. Morgan (Eds.), *Proceedings from the Annual Meeting of the Chicago Linguistic Society 5*, pp. 98–107. Chicago Linguistic Society.
Karttunen, L. and S. Peters (1979). Conventional implicature. In C.-K. Oh and D. A. Dinneen (Eds.), *Syntax and semantics*, Volume 11, pp. 1–56. New York: Academic Press.
Kennedy, C. (2007). Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives. *Linguistics and Philosophy 30*(1), 1–45.
Kennedy, C. and L. McNally (2005). Scale Structure, Degree Modification, and the Semantics of Gradable Predicates. *Language 81*(2), 345–381.
Krifka, M. (2000). Alternatives for aspectual particles: Semantics of *still* and *already*. In L. J. Conathan, J. Good, D. Kavitskaya, A. B. Wulf, and A. C. L. Yu (Eds.), *Proceedings of the 26th Annual Meeting of the Berkeley Linguistics Society*, pp. 401–412.
Lassiter, D. and N. D. Goodman (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. In T. Snider (Ed.), *Proceedings of the 23rd Semantics and Linguistic Theory Conference*, pp. 587–610.
Lombrozo, T. (2016). Explanatory Preferences Shape Learning and Inference. *Trends in Cognitive Sciences 20*(10), 748–759.
Nakanishi, K. and H. Rullmann (2009). Epistemic and concessive interpretations of *at Least*. CLA, Carleton University.
Piñón, C. (2011). The pragmatics of actuality entailment. In *Workshop Aspect and Modality in Lexical Semantics, Stuttgart*, Volume 30.
Rooth, M. E. (1985). *Association with Focus*. Ph. D. thesis, Department of Linguistics, University of Massachusetts, Amherst.
Rooth, M. E. (1992). A theory of focus interpretation. *Natural Language Semantics 1*(1), 75–116.
Rotstein, C. and Y. Winter (2004). Total Adjectives vs. Partial Adjectives: Scale Structure and Higher-Order Modifiers. *Natural Language Semantics 12*(3), 259–288.
Rullmann, H. (1997). Even, polarity, and scope. *Papers in experimental and theoretical linguistics 4*, 40–64.
Rullmann, H. (2007). What does *even* even mean. Ms., University of British Columbia.
Rullmann, H. and K. Nakanishi (2009). More about 'at least'. In *Talk Presented at the MOSAIC Workshop, Ottawa*.
Solt, S. (2011). Notes on the comparison class. In R. Nouwen, R. van Rooij, U. Sauerland, and H.-C. Schmitz (Eds.), *Vagueness in Communication*, Berlin, Heidelberg, pp. 189–206. Springer Berlin Heidelberg.
Stalnaker, R. (2002). Common Ground. *Linguistics and Philosophy 25*(5), 701–721.
Toledo, A. and G. W. Sassoon (2011). Absolute vs. Relative Adjectives - Variance Within vs. Between Individuals. *Semantics and Linguistic Theory 21*, 135–154.
Wilkenfeld, D. A. and T. Lombrozo (2015). Inference to the Best Explanation (IBE) Versus Explaining for the Best Inference (EBI). *Science & Education 24*(9), 1059–1077.