Deriving evaluativity in *even*-comparatives via presupposition accommodation*

Agnes Bi

October 6, 2021

1 Introduction

Discourse participants are constantly negotiating what they take to be in the common ground. From a Stalnakerian point of view, the presupposition φ of an utterance *S* can be accommodated when the addressee recognizes that the speaker intends φ to be in the common ground and, if φ does not contradict the addressee's beliefs, adjusts their own understanding of the common ground accordingly (Stalnaker 2002). It is a well-known and puzzling fact that some presuppositions are easier to accommodate than others. This paper focuses on how the presupposition of scalar particles such as *even* is accommodated in comparative sentences, and how abductive reasoning comes into play in deriving their inferences.

Sentences involving the scalar particle *even* result in evaluative inferences as their default reading when combined with gradable predicates. By *default reading*, I am assuming that the sentences are uttered more or less out of blue, or with minimal additional background information. As (1) shows, the nature of these inferences varies depending on where the focus is placed:

(1) a. Alex is even taller than $[\text{Blake}]_{\text{F}}$. [$\rightsquigarrow$ Both Alex and Blake are **tall**][1]

b. $[\text{Alex}]_{\text{F}}$ is even taller than Blake. [$\rightsquigarrow$ Both Alex and Blake are **short**][2]

*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.

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

[2]Perhaps a better 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 don't reply on *even* to back-associate. Cf. the comparable version of (1-a) *Alex is taller than* *even* $[Blake]_F$. Thanks to Aviv Schoenfeld for this suggestion at CLS poster session.

Positive adjectives such as *tall* by themselves do not produce evaluative inferences in the comparatives. We cannot infer Alex's or Blake's height status from (2) alone:

(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 pitch accent sentence-finally, such as (1-a), 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 the relevant background and review the competing semantic analyses on the market. In section 3, I present the details of my proposal and illustrate why a pragmatic approach is potentially more appealing. Section 4 assesses the proposal by expanding the empirical domain and sketch out possible extensions. And section 5 concludes.

2 Previous analyses

Positive inferences such as that of (1-a) has 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 *p* not being less likely than *q*; (ii) examples where *p* asymmetrically entails *q*, in which case the presupposition of *even* is supposedly satisfied, yet *even p* is still infelicitous; (iii) unexpected sensitivity to standards of comparison. I will leave aside the first two issues for now, and focus on the third in this paper. Greenberg observes that when *even* associates with comparatives formed with relative adjectives, it gives rise to entailments of their positive forms:

(4) a. The blue tool is (even) [stronger than the red tool]$_{\text{F}}$. (Greenberg 2015)

Without *even*: No inference (... both can be weak)

With *even*: Entailment $\rightsquigarrow$ The blue/red tool is strong (# ... both can be weak)

b. Bill is (even) [taller than John]$_F$.

Without *even*: No inference (... both can be short)

With *even*: Entailment $\rightsquigarrow$ Bill/John is tall (# ... both can be 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. Based on this insight, Greenberg and colleagues' work branch into two slightly different analyses, depending on whether the scale of likelihood is retained in the revised meaning of *even*.

2.1 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:

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

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

where *x* is a non-focused entity within the prejacent *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 (5) can easily explain the contrasts between (6-a) and (6-b), which notably poses a problem for the classic analysis:

(6) *Context: John is an accountant, working in a formal office environment.* (Greenberg 2017)

a. John wore a colorful T-shirt to work yesterday, and he even wore [a funny old hat]$_{\mathrm{F}}$.

b. #John wore his usual white shirt to work yesterday, and he even wore [a funny old hat]$_{\mathrm{F}}$.

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 (6-b) is thus predicated to be felicitous, contrary to speakers' intuition. However, if we adopt the proposal in (5), 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's considered unexpected. This violates the Positive Condition and leads to a presupposition failure.

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

(1-a) Alex is even taller than [Blake]$_{\mathrm{F}}$. [$\rightsquigarrow$ Both Alex and Blake are **tall**]

Following Greenberg 2018, assume *even* associates with the Degree Phrase *-er than Blake*[3], and the set of alternatives C is as in (7):

(7) $C_{\text{(1-a)}}$ = { 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 (1-a) is then

(8) **Presupposition of (1-a) per Greenberg 2018**

$\forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_2 \in [\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) > \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake}))] \wedge$
$w_1 \in [\max(\lambda d_1.\text{TALL}(d_1)(\text{Alex})) = \max(\lambda d_2.\text{TALL}(d_2)(\text{Blake}))]] \rightarrow$
$[\mathit{max}(\lambda d_3.\text{TALL}(d_3)(\text{Alex})(w_2))) > \mathit{max}(\lambda d_1.\text{TALL}(d_1)(\text{Alex})(w_1))$ Superlative Condition
$\wedge\, \mathit{max}(\lambda d_1.\text{TALL}(d_1)(\text{Alex})(w_1)) \geqslant \text{Stand}_{\text{TALL}}]]$ Positive Condition

Only when Blake's height is known or fixed in the context, the superlative condition is 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 exactly Blake's height (*q-and-not-p* worlds). The positive condition further requires that Alex's degree of tallness in the worlds where

[3]I assume a phrasal comparison structure here for the ease of illustration. This is not crucial for Greenberg's analysis; phrasal comparison structure goes through in the same fashion.

she is the exactly 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 focus on subject

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

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

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

For the purpose of demonstration, let's assume $\{x|x \in \text{ALT(ALEX)}\}$ 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 (1-b) is thus

(10) **Presupposition of (1-b) per Greenberg 2018 (attempt 1)**

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

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 it is unclear to me why this rescue strategy can be employed, and whether it is always available systematically.

(11) **Presupposition of (1-b) per Greenberg 2018 (attempt 2)**

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

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 that Blake's degree of shortness in the worlds where Alex is taller than her to surpass the standard for shortness. Crucially, the presupposition in (11) 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 (1-a) and Alex in (1-b), 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).

2.2 Adding positive condition to the classic analysis

Daniels and Greenberg 2020 proposes to address the standard sensitivity of *even* in comparatives when the focus is on the source of comparison (cases such as (1-a)) by hard-wiring a positive condition to the meaning of *even*.

(12) **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 left it unclear, let's 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 the individuals in the comparison class have cluster around it. Here is a summary of the assumptions adopted in the paper:

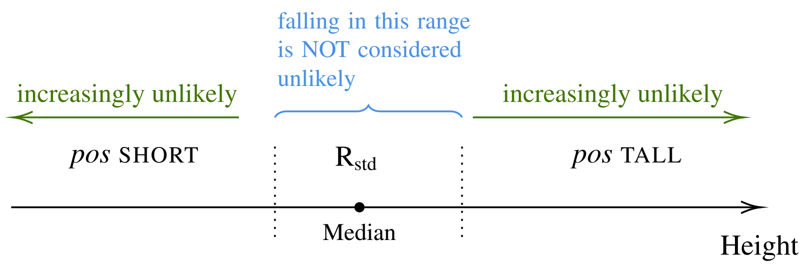
(13) **Daniels and Greenberg 2020's assumptions regarding scales of gradable predicates**

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 $\mathrm{R}_{\mathrm{Std}} >_{\mathrm{likely}}$ having degrees outside $\mathrm{R}_{\mathrm{Std}}$;
d. having degrees within $\mathrm{R}_{\mathrm{Std}}$ cannot be unlikely;
e. variations within $\mathrm{R}_{\mathrm{Std}}$ are allowed, i.e., it is possible to have a total ordering of the individuals fall within $\mathrm{R}_{\mathrm{Std}}$, but these differences are negligible, and thus not considered unlikely.

The following figure demonstrate these assumptions on the scale of height:



For (1-a), 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 $>_{\mathrm{tall}}$ Blake $>_{\mathrm{tall}}$ 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 $>_{\mathrm{tall}}$ Blake
~~*average*~~	~~*average*~~	following assumption (13-e), *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 $>_{\mathrm{tall}}$ Blake
~~*short*~~	~~*average*~~	contradicts Alex $>_{\mathrm{tall}}$ Blake
~~*short*~~	~~*short*~~	since it's not unlikely to be taller than a short individual, *p* is not unlikely

The authors point out that the remaining two cases need further examination: (i) If Alex is tall and Blake is neither tall or 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 the likelihood ranking of the propositions p and q derives 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 (1-a) in a specific context.

(14) *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]$_F$. cf. (1-a)

In this scenario, p = *Alex is taller than Blake* is less likely than q = *Alex is taller than Aaron* simply because that it's more likely for a kid to be taller than someone of their own age than to be taller than someone older. Crucially, it doesn't have to the case that Blake is strictly taller than Aaron in actuality. Perhaps a more clear-cut example is as follows:

(15) *Context: Stephanie was in 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. 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 (12), 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. However, we observe the same inferences with both maximum and minimum standard adjectives, which would have to be explained either by proposing a separate mechanism or by revising the assumptions for Daniels and Greenberg 2020.

(16) *Maximum standard absolute adjective*
   a. District A is even safer than [District B]$_F$. [$\rightsquigarrow$ Both districts are **safe**]
   b. [District A]$_F$ is even safer than District B. [$\rightsquigarrow$ Both districts are **dangerous**]

(17) *Minimum standard absolute adjective*
   a. The study is even dirtier than [the kitchen]$_F$. [$\rightsquigarrow$ Both rooms are **dirty**]
   b. [The study]$_F$ is even dirtier than the kitchen. [$\rightsquigarrow$ Both rooms are **clean**]

2.3 Interim summary

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

Broadly speaking, 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 borne out empirically. In (14), the Smith kids can be tall, short or average height; we don't necessarily infer anything about their absolute height status from the sentence.

Moreover, similar evaluative inferences are seen in other scalar particles such as *at least* and *still*, and it would be ideal to have an uniform analysis, instead of enriching their semantics individually.

(18) 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]

3 A pragmatic approach

Given the aforementioned reasons, this paper takes a pragmatic approach instead, and argues that supplementing the classic (un)-likelihood analysis of *even* with two general pragmatic principles can derive these inferences systematically.

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.

3.1 Proposal

I will illustrate the details of my proposal using the minimal pair in (1) again. 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:

(19) Alternatives of (1-a) = {Alex is taller than $\pi \mid \pi \in$ ALT(BLAKE)}

Alternatives of (1-b) = {π 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. Either this presupposition is already entailed by the common ground, in which case no accommodation is needed and evaluative inference do not necessarily arise, as in (14); or, the addressee learns that this likelihood ranking is entailed by the speaker's intended common ground, and somewhat unexpectedly, commences an abductive reasoning trying to find the most likely justification for this ranking. Discourse participants are not satisfied with simply accommodating the likelihood presupposition on its own; they need to understand why the speaker would presuppose such differences in likelihood exist in the first place. I claim that in the absence of further context, absolute height status is the best and most salient explanation.

Let's adopt a classic formulation of abduction

(20) **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 (1-a), *E* is the fact the speaker's intended common ground entails 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[4], so on and so forth. Among these explanations, H_i Blake is the tallest 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 infers H_i to be true, i.e., Blake is the tallest 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.

(21) **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 at the same time.

For *even*-comparatives such as (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

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

likelihoods is a natural one, considering the *transitive* and *antisymmetric* properties of ordering 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 and so forth).

(22) 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

(23) 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** $\mathcal{L}$

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

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

Note (24) 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 explanation of their corresponding consequent, and given PAC, to accommodate the likelihood rankings, listeners need to accommodate the orderings on the relevant scale as well:

(25) **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 $\mathcal{L}\big(z >_{\text{G}} x\big) < \mathcal{L}\big(z >_{\text{G}} y\big)$ requires that $x >_{\text{G}} y$ is either entailed in the common ground or can be accommodated;

b. Accommodating $\mathcal{L}\big(x >_{\text{G}} z\big) < \mathcal{L}\big(y >_{\text{G}} z\big)$ requires that $y >_{\text{G}} x$ is either entailed in the common ground or can be accommodated.

An immediate predication of this account is that when a presupposition of likelihood ranking(s) is already entailed in the common ground, since there is no strict need for abductive reasoning any more, the evaluative inferences do not necessarily arise. This is borne out empirically, and explains what we observed for (14), repeated below, where discourse participants already have clear reasons to believe that Alex being taller than Blake is less likely.

(14) *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]$_F$. cf. (1-a)

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 don't necessarily infer anything about their height status relative to the standard from the *even*-sentence above.[5]

Going back to the original example where the *even*-comparative is uttered out of blue, the likelihood-to-degree mapping leads to an interim conclusion of where the standard sits relative to other members of the comparison class on the relevant scale. In (1-a) where, unlike (14), there are presumably more than two individuals entering the computation of alternatives, the PAC corollary guarantees that Blake is the tallest within the comparison class.

It is important to note that Blake being the tallest 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 standard. To bridge this gap, let us assume that when an *even*-comparative is uttered more or less out of 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:

(26) **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.

[5]It is interesting to note that there is a secondary, somewhat "hidden" implicature of (14) – 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 rst 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 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's been established in the conversation that Aaron is taller than Blake. It's 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. A rough derivation is as follows: the alternative set here is {Alex is taller than $\pi \mid \pi \in \{\text{Aaron, Blake}\}$}; the relevant G is the property *tall*, $x =$ Blake, $y =$ Aaron, $z =$ Alex. The best explanation for $\mathcal{L}(\text{Alex} >_{\text{TALL}} \text{Blake}) < \mathcal{L}(\text{Alex} >_{\text{TALL}} \text{Aaron})$ is that Blake $>_{\text{TALL}}$ Aaron, and, hence, we assume it to be true. Thanks to Danny Fox (p.c.) for pointing this out.

At its core, all (26) states is for interlocutors to assume normality unless told otherwise.

Now we have all the necessary pieces to give a complete derivation of the positive inference in (1-a) 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 $\{\text{Alex is taller than } \pi \mid \pi \in \text{ALT}(\text{BLAKE})\}$. First, the likelihood-to-degree mapping guarantees that *Blake is the tallest person in the comparison class*. Second, since 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 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 justification for each step is noted above the arrow.

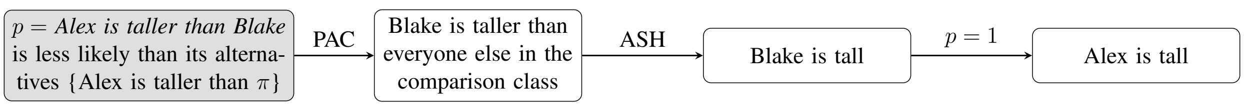


Figure 1: Pragmatic reasoning process for (1-a)

The reversed, negative, inference of (1-b) falls out straightforwardly from the same reasoning process. Let's assume that negative gradable predicates share the same fields as their positive counterparts, but with the opposite relations. Starting with the alternative set $\{\pi \text{ is taller than Blake} \mid \pi \in \text{ALT}(\text{ALEX})\}$, 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 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 (1-a), comprehenders come to conclusion about Blake's height status first, while for (1-b), they do so about Alex's height status first. Therefore, this analysis makes the prediction 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

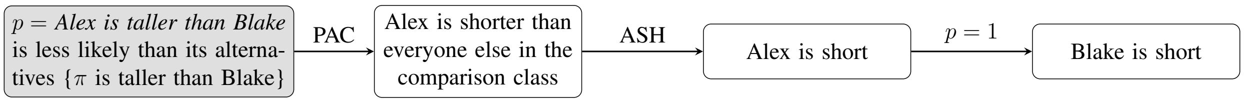


Figure 2: Pragmatic reasoning process for (1-b)

polar questions, we infer the height status of distinct individuals. This seems to be consistent with speakers' judgment:

(27) a. Is Alex even taller than [Blake]$_{\text{F}}$? [$\rightsquigarrow$ Blake is tall]
b. Is [Alex]$_{\text{F}}$ even taller than Blake? [$\rightsquigarrow$ Alex is short]

The intuition behind this proposal is a simple one. 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.

3.2 Testing other types of adjectives

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 (28) uttered out of blue:

(28) a. Pomelos are even cheaper than [pomegranates]$_{\text{F}}$. [$\rightsquigarrow$ Both fruits are **cheap**]
b. [Pomelos]$_{\text{F}}$ are even cheaper than pomegranates. [$\rightsquigarrow$ Both fruits are **expensive**]

For (28-a), the presupposition of *even* requires that the proposition *pomelos are cheaper than pomegranates* is the least likely among its alternatives {pomelos are cheaper than φ | $\varphi \in$ 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 $\mathcal{L}\big(z >_{\text{CHEAP}} x\big) < \mathcal{L}\big(z >_{\text{CHEAP}} y\big)$, 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 (28-b), *p* ‘pomelos are cheaper than pomegranates’ is the least likely among its alternatives $\{\varphi$ 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 $\mathcal{L}\big(x >_{\text{CHEAP}} z\big) < \mathcal{L}\big(y >_{\text{CHEAP}} z\big)$, 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, recall in (16) and (17), repeated below, we observed 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.).

(16) 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**]

(17) 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 (29) or with an overt *for*-phase specifying the comparison class (30), and with respect to entailments (31).

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

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

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

(31) **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 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. A set of assumptions satisfied by relative adjectives, such as the one listed in (13) for Daniels and Greenberg 2020, does not necessarily apply to absolute adjectives. In the interest of theoretical parsimony, these differences should be irrelevant in our discussion here, since we observe the exact same inference pattern. My proposal, relying only on the minimal assumption of the scale being a partial ordering, is appealing in this regard.

3.3 Discussion: Motivating the principles

The Presupposition Accommodation Condition is indirectly supported by the study of *explanations* in both philosophy and psychology (Harman 1965, Lombrozo 2016, Wilkenfeld and Lombrozo 2015). A growing body of work in experimental psychology has revealed that people have a strong, systematic inclination to favor explanations that are simple, broad, and consistent with their prior beliefs. Wilkenfeld and Lombrozo 2015 further argues that "the *process* of seeking, generating, or evaluating explanations itself puts one in a better epistemic position, even when the outcome of the process is not a true explanation" (p. 1060). The act of explaining itself can not only promote abstraction (Williams and Lombrozo 2010) and comparisons between analogous cases (Edwards, Williams, Gentner, and Lombrozo 2019), but also help with recognizing conflicting beliefs (Chi, De Leeuw, Chiu, and Lavancher 1994). Even though this conclusion is drawn from experiments where participants were explicitly asked to give verbal explanations, it is reasonable to assume that the implicit process of seeking and internalizing best explanations shares similar benefits. With regard to the inferences of *even*-comparatives, contrasts on the relevant scale are a simple and probable explanation to why those particular likelihood rankings are assumed by the speaker.

Contextualized in the broader pragmatics literature, the PAC can be considered a mechanism giving rise to one type of *bridging* inferences[7], namely what Clark 1975 categorizes as *reasons, causes, consequences, and concurrences*. 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 that the speaker abides by the Given-New Contract:

(32) **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). For example,

(33) I walked into the room. The chandeliers sparkled brightly. [$\rightsquigarrow$ The room had chandeliers] (Clark 1975, p.171)

Let's refer to the presupposition triggered by the definite description "the chandeliers," i.e., *there existed some chandeliers*, the antecedent of the second sentence. The search for the antecedent was not successful initially, hence the listener filled in the gap by assuming that some chandeliers must be present in the room. 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.

Turing 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[8], 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 alternatives are ranked higher. In other words, all three sentences in (34) assert that *Lydia is three months old*, but with different

[7]Thanks to SuB reviewer #1 for the suggestion.
[8]I would like to thank SuB reviewer #1 for pointing me to this body of work.

"valid" alternatives:

(34) a. Lydia ist [drei]$_F$ monate alt.
Lydia is three months old
Alt = {*Lydia is one month old*, *Lydia is two months old*, *Lydia is three months old*, *Lydia is four months old*, *Lydia is five months old*, ...}

b. Lydia ist **schon** [drei]$_F$ Monate alt.
Lydia is **already** three months old
Alt = {*Lydia is one month old*, *Lydia is two months old*, *Lydia is three months old*}

c. Lydia ist **noch** [drei]$_F$ Monate alt.
Lydia is **still** three months old
'Lydia is only three months old.'
Alt = {*Lydia is three months old*, *Lydia is four months old*, *Lydia is five months old*, ...}
(Krifka 2000, p. 405)

Beaver and Clark 2008 had the similar idea when accounting for mirative function of *only*. *Only*-sentences often carry a sense of unexpectedness or surprise:

(35) London police expected a turnout of 100,000 but **only** 15,000 showed up.
(Beaver and Clark 2008, p. 252, *web example*)

The authors argue that hearer's expectation of something stronger than the prejacent being true is an essential part of the meaning of *only*, and they capture this by presupposing the only 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 two distinct set of data. The Alternative-sampling Hypothesis proposed in this paper is in the same spirit.

4 Conclusion and implications

I have proposed two general cooperative principles governing how rational agents reason about language, which together explain the evaluative inferences that arise in comparative sentences with *even*. This analysis accounts for the optionality of the evaluative inferences, with context (14) versus without (1-a), and straightforwardly derives the puzzling scale reversal in (1) with minimal additional assumptions. In what remains, I will explore how this approach might shed some light on the analyses of concessive *at least* and Mandarin scalar particle *hái*.

4.1 Concessive *at least*

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

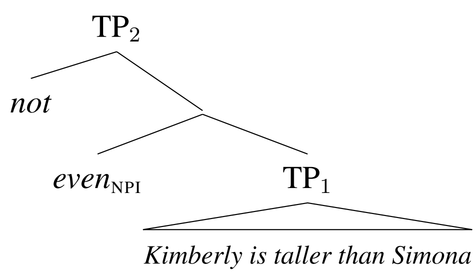
(36) a. Kimberly is even taller than $[\text{Simona}]_F$. [$\rightsquigarrow$ Both are **tall**]
b. Kimberly is <u>not</u> even taller than $[\text{Simona}]_F$. [$\rightsquigarrow$ Both are **short**]

(37) a. $[\text{Kimberly}]_F$ is even taller than Simona. [$\rightsquigarrow$ Both are **short**]
b. $[\text{Kimberly}]_F$ is <u>not</u> even taller than Simona. [$\rightsquigarrow$ Both are **tall**]

Under the NPI theory of *even* (Rooth 1985, Rullmann 1997, Giannakidou 2007)[9], 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 'ordinary' *even* presupposes that 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

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

The evaluative inferences in (36-b) and (37-b) can be straightforwardly derived following the analysis proposed in the last section. Let us assume the LF below for both sentences,

(39)


but importantly, difference in focus placement triggers distinct alternative sets.

For (36-b), the meaning of NPI *even* in (38) presupposes that TP_1 *Kimberly is taller than Simona* is the most likely among its alternatives $\{$Kimberly is taller than $\pi \mid \pi \in \text{ALT}(\textsc{Simona})\}$. 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 is taller than Simona, i.e., Simona is the shortest

[9]NPIs (negative polarity items) are lexical items that are only licensed in negative, downward entailing, or nonveridical 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.

within 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 (37-b) except now the alternative set is $\{\pi$ is taller than Simona $\mid \pi \in$ ALT(KIMBERLY)$\}$, and part (b) of the likelihood-to-degree mapping becomes relevant.

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:

(40) 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:

(41) a. (Rullmann and Nakanishi 2009, p. 17)

Dutch		Japanese	
zelfs	*maar*	*-dake*	*-demo*
even	only	only	even
= NPI 'even'		= 'at least'	

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*.

(42) a. Kimberly is at least taller than [Simona]$_F$. [$\rightsquigarrow$ Both are **short**]

b. [Kimberly]$_F$, at least, is taller then 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 (43-a) or the exact height of Simona in (43-b), 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 did report the judgment above. It would be interesting to investigate why the judgment for (42-a) is considerably clearer than that for (42-b), and whether the same analysis can be extended to account for these inferences.

4.2 Mandarin scalar particle *hái*

The Mandarin scalar particle *hái* is similar to Greek *esto* in that both are sometimes translated as 'even' and sometimes as 'at least' (Giannakidou 2007). Comparable type of reversed evaluative inferences to that of (1-b) versus (1-a) is expressed by the addition of, what is argued by Paul and

Whitman 2008 to be, a cleft construction *shì ... de* instead of change in pitch accent as in English:

(43) a. Jiéyú bǐ Ānqí hái gāo.
J. compared-to A. HAI tall
'Jieyu is even taller than $[\text{Anqi}]_F$.' [$\leadsto$ Both Jieyu and Anqi are **tall**]

b. Jiéyú bǐ Ānqí hái shì gāo de.
J. compared-to A. HAI COP tall DE
'$[\text{Jieyu}]_F$ is even taller than Anqi' [$\leadsto$ Both Jieyu and Anqi are **short**]

I suspect what is happening here is that the cleft construction changes the focus structure of the sentence. As a result, a different set of alternatives is triggered, and enters the calculation of the presupposition of *hái*. As discussed in depth in Lee, Gordon, and Büring 2008, languages differ in the overt cues utilized to cue focus, which can be via morphology, syntactic movement, and/or prosody. I will leave the concrete analysis of the Mandarin case for future research.

In conclusion, even though accounting for evaluative inferences in *even*-comparatives seems straightforward 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.

References

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

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477.

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

Clark, H. H. (1975). Bridging. In *Theoretical issues in natural language processing*.

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

Daniels, M., & Greenberg, Y. (2020). Even and Standards of Comparison.

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

Edwards, B. J., Williams, J. J., Gentner, D., & Lombrozo, T. (2019). Explanation recruits comparison in a category-learning task. *Cognition*, *185*, 21–38.

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

Greenberg, Y. (2015). Even, Comparative Likelihood and Gradability. In *Proceedings of the Amsterdam Colloquium 20*.

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

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

Horn, L. R. (1969). A presuppositional analysis of only and even. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society* (Vol. 5, pp. 98–107). Chicago Linguistic Society.

Karttunen, L., & Peters, S. (1979). Conventional Implicature. *Syntax and semantics*, *11*, 1–56.

Kennedy, C. (2007). Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives. *Linguistics and Philosophy*, *30*(1), 1–45.

Kennedy, C., & McNally, L. (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 *Annual meeting of the Berkeley Linguistics Society* (Vol. 26, pp. 401–412).

Lassiter, D., & Goodman, N. D. (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. *Semantics and Linguistic Theory*, *23*(0), 587–610.

Lee, C., Gordon, M., & Büring, D. (Eds.). (2008). *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*. Springer Science + Business Media B.V.

Lombrozo, T. (2016). Explanatory Preferences Shape Learning and Inference. *Trends in Cognitive Sciences*, *20*(10), 748–759.

Paul, W., & Whitman, J. (2008). Shi … de focus clefts in Mandarin Chinese. *The Linguistic Review*, *25*(3-4), 413–451.

Rooth, M. (1985). *Association with focus*. GLSA, Dept. of Linguistics, University of Massachusetts, Amherst.

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

Rotstein, C., & Winter, Y. (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., & Nakanishi, K. (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, & H.-C. Schmitz (Eds.), *Vagueness in Communication* (pp. 189–206). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer.

Stalnaker, R. (2002). Common Ground. *Linguistics and Philosophy*, *25*(5), 701–721.

Toledo, A., & Sassoon, G. W. (2011). Absolute vs. Relative Adjectives - Variance Within vs. Between Individuals. *Semantics and Linguistic Theory*, *21*(0), 135–154.

Wilkenfeld, D. A., & Lombrozo, T. (2015). Inference to the Best Explanation (IBE) Versus Explaining for the Best Inference (EBI). *Science & Education*, *24*(9), 1059–1077.

Williams, J. J., & Lombrozo, T. (2010). The Role of Explanation in Discovery and Generalization: Evidence From Category Learning. *Cognitive Science*, *34*(5), 776–806.