Cumulative Reading, QUD, and Maximal Informativeness

Linmin Zhang[1,2](✉)

[1] NYU Shanghai, Shanghai, China
[2] NYU-ECNU Institute of Brain and Cognitive Science, Shanghai, China
zhanglinmin@gmail.com,linmin.zhang@nyu.edu
https://sites.google.com/site/zhanglinmin/

Abstract. Motivated by our intuitive interpretation for two kinds of cumulative-reading sentences, this paper argues for a novel **QUD-based view of maximal informativeness**. For a sentence like *Exactly three boys saw exactly five movies* (Brasoveanu 2013), it addresses an underlying QUD like *how high the film consumption is among boys* and provides a most informative answer with mereological maximality. However, for a sentence like *In Guatemala, at most 3% of the population own at least 70% of the land* (Krifka 1999), it addresses rather a QUD like *how skewed wealth distribution is in Guatemala* and provides a most informative answer with the maximality of the ratio between the amount of wealth and its owner population. I implement the analysis of these cumulative-reading sentences within a dynamic semantics framework (à la Bumford 2017). I also compare the current QUD-based view of maximal informativeness with von Fintel et al. (2014)'s entailment-based view and discuss a potentially broader empirical coverage (see also Zhang 2022).

Keywords: Cumulative reading · QUD · Maximal informativeness

1 Introduction

Sentence (1) has a distributive reading (see (1a)) and a cumulative reading (see (1b)). This paper focuses on its **cumulative reading**.

(1) Exactly three boys saw exactly five movies.

a. There are in total 3 boys, and for each atomic boy, there are in total 5 movies such that he saw them. **Distributive reading**
⇝ In total, there are 15 movie-seeing events, and the cardinality of distinct movies involved is between 5 and 15.

This project was financially supported by the Shanghai Municipal Education Commission (the Program for Eastern Young Scholars at Shanghai Institutions of Higher Lerning, PI: L.Z.). I thank the organizers, anonymous reviewers, and audience of LENLS 19. Errors are mine.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
D. Bekki et al. (Eds.): LENLS 2019, LNCS 14213, pp. 1–17, 2023.
https://doi.org/10.1007/978-3-031-43977-3_1

b. The cardinality of all boys who saw any movies is 3, and the cardinality of all movies seen by any boys is 5. **Cumulative reading**
$\rightsquigarrow$ In total, the cardinality of distinct movies involved is 5, and there are between 5 and 15 movie-seeing events.

Our intuition for the cumulative reading of (1) crucially involves the notion of **maximality**. As described in (1b), the two modified numerals (see the underlined parts of (1)) denote and count the **totality** of boys who saw any movies and the **totality** of movies seen by any boys.

In Brasoveanu (2013)'s compositional analysis of the cumulative reading of (1), two mereology-based maximality operators are applied simultaneously (at the sentential level) to derive the truth condition that matches our intuition.

In this paper, I further investigate the nature and source of this maximality. In particular, I follow Krifka (1999) to show that there are natural language cumulative-reading sentences that cannot be naturally interpreted with mereological maximality.

In a nutshell, I propose that (i) although the cumulative reading of (1) involves multiple modified numerals, it actually does not involve multiple independent maximality operators, but only one, and (ii) this maximality operator is not necessarily mereology-based, but rather informativeness-based, with regard to the resolution of a contextually salient degree QUD (Question under discussion). Thus Brasoveanu (2013)'s analysis for (1b) can be considered a special case within a more generalized theory on maximal informativeness.

The rest of the paper is organized as follows. Section 2 presents Brasoveanu (2013)'s mereological-maximality-based analysis of cumulative-reading sentences like (1). Then Sect. 3 presents Krifka (1999)'s discussion on a case that challenges a direct extension of Brasoveanu (2013)'s analysis. In Sect. 4, I propose to adopt the notion of QUD-based maximality of informativeness and show how this new notion of maximality provides a unified account for the data addressed by Brasoveanu (2013) and Krifka (1999). Section 5 compares the current QUD-based view with von Fintel et al. (2014)'s entailment-based view on maximality of informativeness. Section 6 further shows a wider empirical coverage for the notion of QUD-based maximality of informativeness. Section 7 concludes.

2 Brasoveanu (2013)'s Analysis of Cumulative Reading

Cumulative-reading sentences involve modified numerals, which bring **maximality** (see e.g., Szabolcsi 1997, Krifka 1999, de Swart 1999, Umbach 2006).

The contrast in (2) shows that compared to **bare numerals** (here *two dogs*), **modified numerals** (here *at least two dogs*) convey maximality, as evidenced by the infelicity of the continuation *perhaps she fed more* in (2b). Thus, the semantics of *two* in (2a) is **existential**, but the semantics of *at least two* in (2b) is **maximal**, indicating the cardinality of the **totality** of dogs fed by Mary.

(2) a. Mary fed two dogs. They are cute. Perhaps she fed more.
b. Mary fed at least two dogs. They are nice. #Perhaps she fed more.

According to Brasoveanu (2013), the semantics of the cumulative reading of (1) involves the **simultaneous** application of two maximality operators.

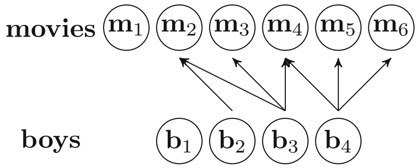


Fig. 1. The genuine **cumulative** reading of (1) is **true** in this context.

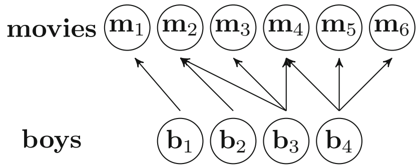


Fig. 2. The genuine **cumulative** reading of (1) is **false** in this context.

(3) Exactly three^u boys saw exactly five$^\nu$ movies. (= (1))
Cumulative reading of (1): (Brasoveanu 2013)

$$\underbrace{\sigma x \sigma y[\textsc{boy}(x) \wedge \textsc{movie}(y) \wedge \textsc{see}(x,y)]}_{\text{the mereologically maximal } x \text{ and } y \text{ satisfying these restrictions}} \wedge \underbrace{|y| = 5 \wedge |x| = 3}_{\text{cardinality tests}}$$

(σ: maximality operator; for notation simplicity, cumulative closure is assumed.)

As sketched out in (3), Brasoveanu (2013) casts his analysis in dynamic semantics. The semantic contribution of modified numerals is two-fold. They first introduce plural discourse referents (drefs), x and y (assigned to u and ν respectively). Then after restrictions like $\textsc{boy}(x)$, $\textsc{movie}(y)$, and $\textsc{see}(x, y)$ are applied onto these drefs, modified numerals contribute maximality tests and cardinality tests. As shown in (3), two maximality operators σ are applied simultaneously to x and y at the global/sentential level, picking out the mereologically maximal x and y that satisfy all the relevant restrictions. Finally, these mereologically maximal x and y are checked for their cardinality, so that eventually the sentence addresses the cardinality of all the boys who saw any movies (which is 3) and the cardinality of all the movies seen by any boys (which is 5).

Crucially, the genuine cumulative reading characterized in (3) is distinct from the non-attested pseudo-cumulative reading shown in (4), where *exactly three boys* takes a wider scope than *exactly five movies*:

(4) **Unattested pseudo-cumulative reading of (1)**:

$$\underbrace{\sigma x[\textsc{boy}(x) \wedge \underbrace{\sigma y[\textsc{movie}(y) \wedge \textsc{see}(x,y)]}_{\text{the mereologically maximal } y} \wedge |y| = 5]}_{\text{the mereologically maximal } x} \wedge |x| = 3$$

i.e., The maximal plural individual x satisfying the restrictions (i.e., atomic members of x are boys and each of them saw some movies, and x saw a total of 5 movies between them) has the cardinality of 3.
$\rightsquigarrow$ **True** for Fig. 2! (see $b_2 \oplus b_3 \oplus b_4$ and $b_1 \oplus b_2 \oplus b_4$, and there is no larger boy-sum satisfying these restrictions.)

The analysis shown in (4) can be ruled out by the contrast between our intuitive judgments: sentence (1) is judged true under the context shown in Fig. 1 but false under the context shown in Fig. 2. However, the truth condition characterized in (4) is actually true under the context shown in Fig. 2, where boy-sums $b_2 \oplus b_3 \oplus b_4$ and $b_1 \oplus b_2 \oplus b_4$ each saw a total of 5 movies, and there are no larger boy-sums such that they saw in total 5 movies between them. Therefore, as concluded by Brasoveanu (2013), only (3), but not (4), captures our intuitive interpretation of sentence (1). In other words, our intuitive interpretation for the cumulative reading of (1) (see (3)) involves no scope-taking between the two modified numerals (see also Krifka 1999, Charlow 2017 for more discussion).

Although Brasoveanu (2013) and relevant discussions in existing literature have shown that the reading of (4) is empirically non-attested and needs to be ruled out, they do not explain **why** (4) is not attested. In this sense, the simultaneity of applying two maximality operators seems a stipulation.

3 A Challenging Case Discussed by Krifka (1999)

Krifka (1999) uses sentence (5) to address his observations on cumulative reading.

(5) In Guatemala, <u>at most 3%</u> of the population own <u>at least 70%</u> of the land. (≈ (13) and (27) of Krifka 1999)

First, the intuitively most natural interpretation of (5) also indicates that there is no scope-taking between the two modified numerals here:

> 'The problem cases discussed here clearly require a representation in which NPs are not scoped with respect to each other. Rather, they ask for an interpretation strategy in which all the NPs in a sentence are somehow interpreted in parallel, which is not compatible with our usual conception of the syntax/semantics interface which enforces a linear structure in which one NP takes scope over another.' (Krifka 1999)

Then Krifka 1999 further points out that the simultaneous mereology-based maximization strategy that works for data like (1) does not work for (5):

> 'Under the simplifying (and wrong) assumption that foreigners do not own land in Guatemala, and all the land of Guatemala is owned by someone, this strategy would lead us to select the alternative *In Guatemala, 100 percent of the population own 100% of the land*, which clearly is not the most informative one among the alternatives – as a matter of fact, it is pretty uninformative. We cannot blame this on the fact that the NPs in (27) (i.e., (5) in the current paper) refer to percentages, as we could equally well express a similar statistical generalization with the following sentence (assume that Guatemala has 10 million inhabitants and has an area of 100,000 square kilometers):
>
> (28) In Guatemala, 300,000 inhabitants own 70,000 square kilometers of land.

> Again, the alternative *In Guatemala, 10 million inhabitants own 100,000 square kilometers of land* would be uninformative, under the background assumptions given.
> What is peculiar with sentences like (27) is that they want to give information about the bias of a statistical distribution. One conventionalized way of expressing particularly biased distributions is to select a small set among one dimension that is related to a large set of the other dimension. Obviously, to characterize the distribution correctly, one should try to decrease the first set, and increase the second. In terms of informativity of propositions, if (27) is true, then there will be alternative true sentences of the form *In Guatemala*, n *percent of the population own* m *percent of the land*, where n is greater than *three*, and m is smaller than *seventy*. But these alternatives will not entail (27), and they will give a less accurate picture of the skewing of the land distribution.' (krifka 1999)

In short, Krifka (1999)'s discussion on (5) suggests that in accounting for cumulative-reading sentences, (i) a direct application of simultaneous mereology-based maximization strategy does not always work, and (ii) what kind of concern interlocutors aim to address via the use of a cumulative-reading sentence matters for sentence interpretation, and in particular, the interpretation of the interplay between modified numerals.

4 Proposal: QUD-Based Maximal Informativeness

As suggested by Krifka (1999), QUD should matter in our intuitive interpretation of cumulative-reading sentences. Based on this idea, I start with an informal discussion on the underlying QUD in interpreting cumulative-reading sentences (Sect. 4.1). Then I propose a QUD-based view on maximality of informativeness (Sect. 4.2) and develop a compositional analysis for cumulative-reading sentences like (1) and (5) within a dynamic semantics framework (Sect. 4.3), à la Bumford (2017) and in the same spirit as Brasoveanu (2013).

4.1 Cumulative-Reading Sentences and Their Underlying QUD

Here I first show that numerals or measure phrases provide quantity/measurement information, but quantity/measurement information alone does not determine how we interpret an uttered sentence and reason about its informativeness. The same sentences (e.g., (6) and (7)) can lead to different patterns of meaning inference, depending on a potentially implicit **degree QUD**.

In (6), the measurement information provided by *7 o'clock* directly addresses *what time it is* (see (6a)). However, it is not (6a), but rather an underlying **degree question**, that determines whether (6) is interpreted as *it's* <u>*as late as*</u> *7 o'clock* (≈ <u>already</u> 7 o'clock) or *it's* <u>*as early as*</u> *7 o'clock* (≈ <u>only</u> 7 o'clock).

If the underlying QUD is *how late it is* (see (6b)), then (6) is interpreted as *it's* <u>*as late as*</u> *7 o'clock*, conveying a stronger meaning than *it's 6/5/... o'clock* by indicating a higher degree of **lateness**. Thus, to resolve *how late it is*, we consider a temporal scale from earlier to later time points, and **higher informativeness** correlates with **later** time points, i.e., the **increase** of numbers.

On the other hand, if the underlying QUD is *how early it is* (see (6c)), then (6) is interpreted as *it's as early as 7 o'clock*, conveying a stronger meaning than *it's 8/9/... o'clock* by indicating a higher degree of **earliness**. Thus, to resolve *how early it is*, we consider rather a temporal scale from later to earlier time points, and **higher informativeness** correlates with **earlier** time points, i.e., the **decrease** of numbers.

(6) It's 7 o'clock.

a. What time is it? **Neutral**: no evaluativity
b. QUD: How late is it? (6) $\rightsquigarrow$ It's as late as 7 o'clock
In addressing (6b), *It's as late as 7:00* $>_{\text{info}}$ *It's as late as 6:00*
c. QUD: How early is it? (6) $\rightsquigarrow$ It's as early as 7 o'clock
In addressing (6c), *It's as early as 7:00* $>_{\text{info}}$ *It's as early as 8:00*

Similarly, along a scale of length, we intuitively feel that *John is 5 feet tall* is stronger than *John is 4 feet tall*. This intuition is actually based on the degree QUD – *How tall is John*. However, it is not guaranteed that measurement sentences containing a higher number are always more informative. Depending on whether the underlying degree QUD is (7b) or (7c), (7) can be interpreted as *John is as tall as 5 feet* and more informative than an alternative sentence with a smaller number, or (7) can be interpreted as *John is as short as 5 feet* and more informative than an alternative sentence with a larger number.[1]

(7) John measures 5 feet.

a. How many feet does John measure? **Neutral**: no evaluativity
b. QUD: How tall is John? (7) $\rightsquigarrow$ He is as tall as 5 feet
In addressing (7b), *John is as tall as 5′* $>_{\text{info}}$ *John is as tall as 4′*
c. QUD: How short is John? (7) $\rightsquigarrow$ He is as short as 5 feet
In addressing (7c) *John is as short as 5′* $>_{\text{info}}$ *John is as short as 6′*

Therefore, as illustrated by (6) and (7), in the interpretation of sentences containing numerals, it is not always the case that the use of larger numbers leads to higher level of informativeness. Rather, the inference on informativeness

[1] Degree questions like *how tall is John* are more default (i.e., less marked) than *how short is John*. Thus, we naturally feel that *John measures 5 feet* (or *John is 5 feet tall*) is stronger (i.e., more informative) than *John measures 4 feet* (or *John is 4 feet tall*). However, I make a distinction between **being more informative** and **entailment** and avoid the term 'entailment' here. As shown in (i), (ia) is stronger than but does not directly entail (ib). See also Sect. 5 for more discussion.

(i) a. John is above 6 feet tall. $\lambda w.\text{HEIGHT}(\text{JOHN})(w) \subseteq [6', +\infty)$
b. John is between 4 and 5 feet tall. $\lambda w.\text{HEIGHT}(\text{JOHN})(w) \subseteq [4', 5']$

.

hinges on (i) an underlying degree QUD (along with the direction of the scale associated with the degree QUD) and (ii) how numerals are used to resolve the degree QUD. Sometimes the use of smaller measurement numbers leads to higher informativeness in resolving degree QUDs (e.g., (6c) and (7c)).

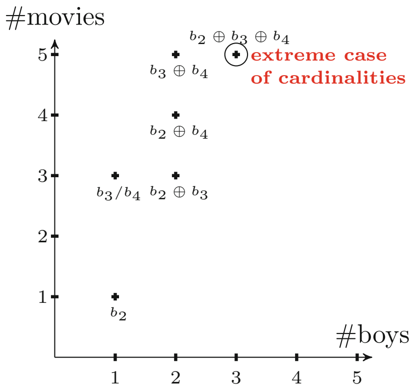


Fig. 3. QUD: How much is the overall film consumption among boys? The cardinalities of some boy-sums and movie-sums in the context of Fig. 1 are plotted as dots. The extreme case that addresses the degree QUD in the most informative way is represented by the right-uppermost dot, i.e., the one corresponding to the boy-sum $b_2 \oplus b_3 \oplus b_4$ and the 5 movies they saw between them.

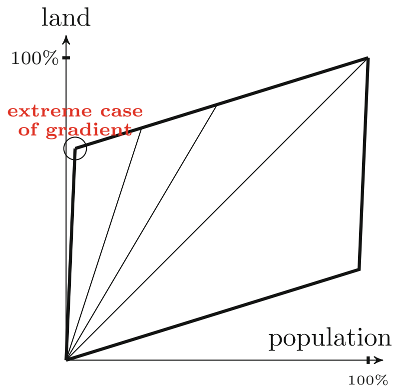


Fig. 4. QUD: How skewed is wealth distribution? The plotting of the percentages of the population and their owned land should form a parallelogram-like area. The extreme case that addresses the degree QUD in the most informative way is represented by the left-uppermost corner, which means that 3% of the population own 70% of the land.

The above observation can be extended to cumulative-reading sentences that contain multiple numerals: the interpretation of a sentence and our inference on its informativeness depend on its underlying degree QUD.

In particular, when multiple numerals are used together to address a degree QUD, their interplay brings new patterns for connecting numbers and meaning inference on informativeness. Higher informativeness does not correlate with the increase or decrease of a **single number**, but **an interplay among numbers**.

According to our intuition, the cumulative-reading of sentence (1) addresses and can be a felicitous answer to QUDs like (8a) or (8b) , but it does not address QUDs like (8c) or (8d).[2] Therefore, as illustrated in Fig. 3, **higher informativeness** correlates with the **increase along both the dimensions** of boy-cardinality and movie-cardinality, and the right-uppermost dot on this two-dimensional coordinate plane represents maximal informativeness. I.e., maximal informativeness amounts to simultaneous mereology-based maximality.

[2] (8c) and (8d) do not even sound like natural questions for some native speakers.

(8) Exactly three boys saw exactly five movies. (= (1))

a. QUD: How many boys saw how many movies?
b. QUD: How much is the overall film consumption among boys?
c. $\not\rightsquigarrow$??How many boys saw exactly five movies (between them)?
d. $\not\rightsquigarrow$??How many movies did exactly three boys see (between them)?

On the other hand, as pointed out by Krifka (1999), the cumulative-reading of sentence (5) addresses and can be a felicitous answer to degree QUDs like (9a), but it does not address QUDs like (9b) (cf. (8a) as a felicitous QUD for (8)). Therefore, as illustrated in Fig. 4, the plotting of the percentage of the population and their entire owned land forms a parallelogram-like area, and **higher informativeness** correlates with a **higher ratio** between the quantity of owned land and its owner population. In other words, higher informativeness correlates simultaneously with the **decrease** along the dimension of population and the **increase** along the dimension of land quantity. It is the left-uppermost corner of this parallelogram-like area that represents maximal informativeness. In this case, obviously, maximal informativeness is not mereology-based.

(9) In Guatemala, at most 3% of the population own at least 70% of the land. (= (5))

a. QUD: How skewed is wealth distribution in Guatemala?
b. $\not\rightsquigarrow$ How many people own how much land in Guatemala?

In brief, although the interpretation of both types of cumulative-reading sentences is based on maximal informativeness, they are different with regard to how maximal informativeness is computed from numbers, and crucially, this computation is driven by an underlying degree QUD.

Further evidence comes from the monotonicity of numerals used in these cumulative-reading sentences. In the case represented by (1)/(8), the two numerals cannot have opposite monotonicity, while in the case represented by (5)/(9), the presence of two numerals with opposite monotonicity is perfectly natural (e.g., in (5)/(9), the use of a downward-entailing expression, *at most 3%*, along with the use of an upward-entailing expression, *at least 70%*). Evidently, in the former case, the two numerals contribute to the informativeness of a sentence in a parallel way, while in the latter case, the two numerals contribute to the informativeness of a sentence in opposite ways.

It is worth mentioning that **multi-head comparatives** (see von Stechow 1984, Hendriks and De Hoop 2001) also provide empirical support for (i) a degree-QUD-based interpretation and informativeness inference as well as (ii) the connection between QUDs and the pattern of monotonicity.

As illustrated in (10)–(12), the underlying QUD determines how the changes of quantity/measurement contribute to sentence interpretation.

In contrast, (13) sounds degraded because with the use of *fewer dogs* and *more rats*, the sentence fails to suggest a QUD that it can felicitously address: (i) the evaluation in terms of the quantity and quality of preys and (ii) the quantity of dogs as successful predators are at odd with each other in conveying coherent meaning.

(10) Less land produces more corn than ever before. (von Stechow 1984, Hendriks and De Hoop 2001)
QUD: How is the productivity rate increased?
⇝ Correlating with the **decrease** of input and the **increase** of output

(11) Nowadays, more goods are carried faster. (Hendriks and De Hoop 2001)
QUD: How is the efficiency of transportation increased?
⇝ Correlating with the **increase** of both amount and speed

(12) More dogs ate more rats than cats ate mice. (von Stechow 1984, Hendriks and De Hoop 2001)
QUD: How are dogs more successful predators than cats?
⇝ Comparison along two dimensions address the QUD in a parallel way

(13) *Fewer dogs ate more rats than cats ate mice. (Hendriks and De Hoop 2001)

4.2 A QUD-Based Maximality Operator

Based on the informal discussion in Sect. 4.1, I propose a **QUD-based maximality operator** and implement it within a dynamic semantics framework:

(14) $\mathbf{M}_{u_1,u_2,\ldots} \stackrel{\text{def}}{=}$
$\lambda m.\lambda g.\,\{h \in m(g) | \neg\exists h' \in m(g)\,.\, G_{\text{QUD}}(h'(u_1, u_2, \ldots)) >_{\text{info}} G_{\text{QUD}}(h(u_1, u_2, \ldots))\}$
(Type of m: $g \rightarrow \{g\}$; Type of $\mathbf{M}$: $(g \rightarrow \{g\}) \rightarrow (g \rightarrow \{g\})$)

As shown in (14), I assume meaning derivation to be a series of updates from an information state to another, and an information state m (of type $g \rightarrow \{g\}$) is represented as a function from an input assignment function to an output set of assignment functions (see also Bumford 2017).

The QUD-based maximality operator $\mathbf{M}_{u_1,u_2,\ldots}$ works like a filter on information states. With the application of $\mathbf{M}_{u_1,u_2,\ldots}$, the discourse referents (drefs, which are assigned to $u_1, u_2, \ldots$) that lead to the maximal informativeness in resolving a QUD will be selected out.

More specifically, the definition of $\mathbf{M}_{u_1,u_2,\ldots}$ includes an operator G_{QUD}, which, when applied on drefs, returns a value indicating informativeness. This informativeness amounts to a measurement in addressing a contextually salient degree QUD: e.g., in the case of (8) (see Fig. 3), *how much the overall film consumption is among boys*; in the case of (9) (see Fig. 4), *how skewed wealth distribution is in Guatemala*. In this sense, G_{QUD} can be considered a measure function.

4.3 Analyzing Cumulative-Reading Sentences

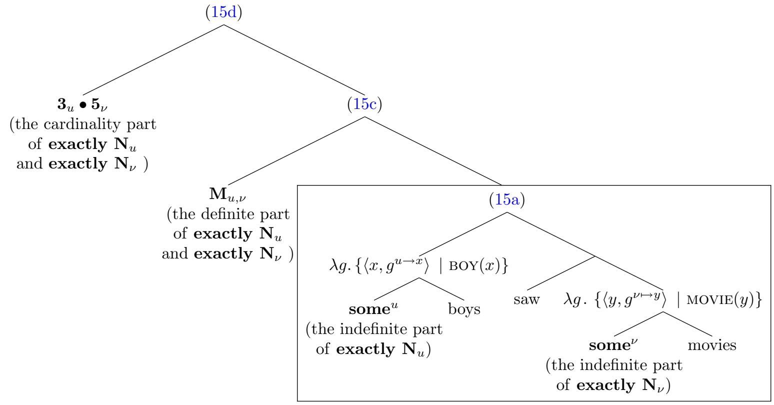
The step-by-step semantic derivation for the core example (1) is shown in (15).

(15a) first shows the introduction of plural drefs and relevant restrictions.

Given that this sentence is interpreted with a contextually salient QUD like *how high film consumption is among boys* (see (8) and Fig. 3), higher informativeness amounts to higher degree of consumption level (e.g., with $d_1 > d_2$, *the* *consumption level is* d_1*-high* is more informative than *the consumption level is* d_2*-high*). Thus the measurement of informativeness amounts to the measurement of cardinalities of both plural drefs (see (15b)).

Maximal informativeness is achieved when the mereologically maximal drefs (i.e., $b_2 \oplus b_3 \oplus b_4$ and $m_2 \oplus m_3 \oplus m_4 \oplus m_5 \oplus m_6$ in Fig. 1) are assigned (see (15c)).

(15) Exactly three^u boys saw exactly five$^\nu$ movies. (= (1))



a. $p = [\![\text{some}^u \text{ boys saw some}^\nu \text{ movies}]\!] =$
$$\lambda g.\left\{ g^{u\mapsto x}_{\nu\mapsto y} \,\middle|\, \textsc{movie}(y), \textsc{boy}(x), \textsc{saw}(x,y) \right\}$$
(i.e., the introduction of plural drefs x and y and restrictions)

b. $G_{\text{QUD}} = \lambda x.\lambda y.|x| + |y|$
(i.e., Based on the QUD, maximizing informativeness amounts to simultaneously maximizing x and y.)

c. $\mathbf{M}_{u,\nu}(p) =$
$$\lambda g.\left\{ g^{u\mapsto x}_{\nu\mapsto y} \,\middle|\, \begin{array}{l} y = \iota y.[\textsc{movie}(y) \wedge \underbrace{\exists x[\textsc{boy}(x) \wedge \textsc{saw}(x,y)]}_{\text{some boys saw } y} \\ \wedge \underbrace{\forall y' \neq y[\textsc{movie}(y') \wedge \exists x[\textsc{boy}(x) \wedge \textsc{saw}(x,y')] \rightarrow y' \sqsubset y]}_{y \text{ is mereologically maximal}}\,] \\ x = \iota x.[\textsc{boy}(x) \wedge \underbrace{\exists y[\textsc{movie}(y) \wedge \textsc{saw}(x,y)]}_{x \text{ saw some movies}} \\ \wedge \underbrace{\forall x' \neq x[\textsc{boy}(x) \wedge \exists y[\textsc{movie}(y) \wedge \textsc{saw}(x,y)] \rightarrow x' \sqsubset x]}_{x \text{ is mereologically maximal}}\,] \end{array} \right\}$$
(i.e., the drefs x and y that lead to maximal informativeness are picked out $\rightsquigarrow$ mereologically maximal x and y are picked out.)

d. $[\![(1)]\!] = [\![\text{exact } 3^u \text{ boys saw exactly } 5^\nu \text{ movies}]\!] = \mathbf{M}_{u,\nu}(p)$, if $\begin{array}{l}|x| = 3, \\ |y| = 5\end{array} =$
$$\lambda g.\left\{ g^{u\mapsto x}_{\nu\mapsto y} \,\middle|\, \begin{array}{l} y = \iota y.[\textsc{movie}(y) \wedge \exists x[\textsc{boy}(x) \wedge \textsc{saw}(x,y)] \\ \wedge \forall y' \neq y[\textsc{movie}(y') \wedge \exists x[\textsc{boy}(x) \wedge \textsc{saw}(x,y')] \rightarrow y' \sqsubset y]\,] \\ x = \iota x.[\textsc{boy}(x) \wedge \exists y[\textsc{movie}(y) \wedge \textsc{saw}(x,y)] \\ \wedge \forall x' \neq x[\textsc{boy}(x) \wedge \exists y[\textsc{movie}(y) \wedge \textsc{saw}(x,y)] \rightarrow x' \sqsubset x]\,] \end{array} \right\},$$
if $|x| = 3$ and $|y| = 5$
(i.e., the cardinalities of mereologically maximal x and y are checked.)

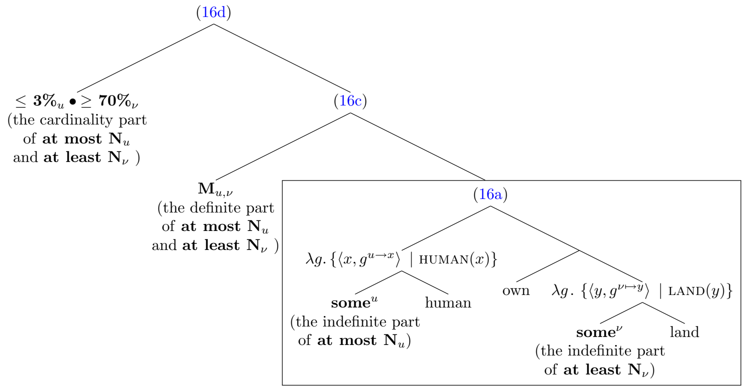
The step-by-step semantic derivation of the core example (5) is shown in (16). The crucial difference between the analysis in (15) vs. (16) consists in their QUD, i.e., G_{QUD}, as reflected in (15b) vs. (16b).

Given that (5) is interpreted with a QUD like *how skewed wealth distribution is in Guatemala* (see (9) and Fig. 4), higher informativeness amounts to higher degree of skewedness. Thus the measurement of informativeness amounts to the ratio between the quantity of drefs (see (16b)).

Maximal informativeness is achieved when the quantity of a dref y satisfying $\textsc{land}(y) \wedge \textsc{own}(x, y)$ divided by the quantity of a dref x satisfying $\textsc{human}(x) \wedge \textsc{own}(x, y)$ yields the maximal ratio/quotient (see (16c)).

Finally, modified numerals *at most 3%* and *at least 70%* impose cardinality tests on the drefs selected out from the step in (16c) (see (16d)).

(16) At most $3\%^u$ of the population own at least $70\%^\nu$ of the land. (=(5))



a. $p = [\![\text{some}^u \text{ population own some}^\nu \text{ land}]\!] =$
$\lambda g.\left\{ g^{\nu\mapsto y}_{u\mapsto x} \;\middle|\; \textsc{land}(y), \textsc{human}(x), \textsc{own}(x, y) \right\}$
(i.e., the introduction of drefs x and y and restrictions)

b. $G_{\text{QUD}} = \lambda x.\lambda y.|y| \div |x|$
(i.e., Based on the QUD, maximizing informativeness amounts to maximizing the ratio between the quantity of y and x.)

c. $\mathbf{M}_{u,\nu}(p) =$

$$\lambda g.\left\{ g^{\nu\mapsto y}_{u\mapsto x} \;\middle|\; \begin{array}{l} \langle x, y\rangle = \langle \iota x, \iota y\rangle \text{ such that} \\ \textsc{land}(y) \wedge \textsc{human}(x) \wedge \textsc{own}(x, y) \\ \wedge \underbrace{\neg\exists y' \sqsupset y[\textsc{land}(y') \wedge \textsc{own}(x, y')]}_{y \text{ is the maximal land owned by } x} \\ \wedge \underbrace{\neg\exists x' \sqsupset x[\textsc{human}(x') \wedge \textsc{own}(x', y)]}_{x \text{ is the maximal owner of } y} \\ \wedge\forall x''\forall y''[\textsc{land}(y'') \wedge \textsc{human}(x'') \wedge \textsc{own}(x'', y'') \\ \wedge \underbrace{\neg\exists y''' \sqsupset y''[\textsc{land}(y''') \wedge \textsc{own}(x'', y''')]}_{y'' \text{ is the maximal land owned by } x''} \\ \wedge \underbrace{\neg\exists x''' \sqsupset x''[\textsc{human}(x''') \wedge \textsc{own}(x''', y'')]}_{x'' \text{ is the maximal owner of } y''} \rightarrow \frac{|y|}{|x|} \geq \frac{|y''|}{|x''|}] \end{array} \right\}$$

(i.e., the drefs x and y that lead to maximal $\frac{|y|}{|x|}$ are picked out.)

d. $[\![(5)]\!] =$
$[\![$at most $3\%^{u}$ of the population own at least $70\%^{\nu}$of the land$]\!] =$
$\mathbf{M}_{u,\nu}(p)$, if $|x| \subseteq (0, 3\%], |y| \subseteq [70\%, 1]$
(i.e., the cardinalities of selected x and y are checked.)

Here I would also like to address an issue raised by an anonymous reviewer: is it valid to use QUD in the derivation of (truth-conditional) meaning of a sentence? Works like Grindrod and Borg (2019) point out that the framework of QUD is pragmatic, mainly accounting for phenomena like the use of prosodic focus in question-answer congruence, and further extension to account for truth-conditional meaning is illegitimate.

As sketched in the following examples (17)–(19), modified numerals actually parallel exactly with items that bear prosodic focus: their interpretation all involve (i) a certain QUD, (ii) the application of a maximality operator that results in maximal informativeness (in addressing the QUD), and (iii) a further post-suppositional test (on cardinality or identity/part-whole relationship). Thus the current proposed treatment of sentences containing modified numerals is actually compatible with, not against, the view of Grindrod and Borg (2019) (see also Krifka 1999; Zhang 2023 for relevant discussions).

(17) Mary read $[\textit{Sandman}]^{u}_{F}$. QUD: What books did Mary read?
$\rightsquigarrow$ A maximality operator picks out the mereologically maximal dref X, such that X is a book-sum read by Mary, and ***Sandman*** works as a post-suppositional test, checking whether X is (or includes) ***Sandman***.

(18) $[\text{Mary}]^{u}_{F}$ read *Sandman*. QUD: Who read *Sandman*?
$\rightsquigarrow$ A maximality operator picks out the mereologically maximal dref X, such that X is a human-sum that read *Sandman*, and **Mary** works as a post-suppositional test, checking whether X is (or includes) **Mary**.

(19) Mary fed at least two^{u} dogs. (= (2b))
QUD: how many dogs did Mary feed?
$\rightsquigarrow$ A maximality operator picks out the mereologically maximal dref X, such that X is a dog-sum fed by Mary, and **at least two** checks the cardinality of this maximal X, whether $|X| \geq 2$.

5 Discussion: Comparison with von Fintel et al. (2014)

Under the current analysis, it is a contextually salient degree QUD (i.e., what interlocutors care about, their ultimate motivation behind their utterance, see Roberts 2012) that determines how informativeness is actually measured (see the implementation of G_{QUD} in (15b) vs. (16b)). This degree-QUD-based informativeness measurement, G_{QUD}, further determines how the maximality operator $\mathbf{M}_{u_1,u_2,\ldots}$ filters on drefs and how (modified) numerals affect meaning inference.

The notion of **degree-QUD-based** maximality of informativeness proposed here is in the same spirit as but more generalized than the **entailment-based** one proposed by von Fintel et al. (2014) (which primarily aims to account for the interpretation of *the*; see also Schlenker 2012). According to von Fintel et al. (2014), informativeness ordering is based on entailment relation (see (20)).

(20) von Fintel et al. (2014)'s notion of informativeness ordering:
For all x, y of type α and property ϕ of type $\langle s, \langle \alpha, t \rangle \rangle$, $x \geq_{\phi} y$ iff $\lambda w.\phi(w)(x)$ entails $\lambda w.\phi(w)(y)$. (von Fintel et al. (2014): (3b))

Thus as shown in (21), depending on the monotonicity of properties, maximal informativeness corresponds to maximum or minimum values.

(21) a. For **upward monotone** properties (e.g., λd. Miranda is d tall), **maximal** informativeness means **maximum** values:
e.g., *Miranda is* 1.65 *m tall* entails *Miranda is* 1.60 *m tall.*
b. For **downward monotone** properties,
maximal informativeness means **minimum** values:
e.g., given that $m > n$, n *walnuts are sufficient to make a pan of baklava* entails m *walnuts are sufficient to make a pan of baklava.*

Compared to von Fintel et al. (2014), the notion of QUD-based maximal informativeness developed in the current paper is more generalized in two aspects.

First, the current QUD-based view on maximality of informativeness can be easily extended from dealing with a single value to a combination of values.

As shown in Sect. 4, in cumulative-reading sentences where multiple numerals are involved, maximal informativeness does not directly correspond to whether the uttered numbers are considered maximum or minimum values. In example (5), as observed by Krifka (1999), each of the numerals (i.e., *at most 3%* and *at least 70%*) alone cannot be maximum or minimum values. It is how the combination of these uttered numbers contributes to resolve an implicit, underlying QUD that leads to the achievement of maximal informativeness.

Second, and more importantly, the current degree-QUD-based view on maximality of informativeness can overcome the issue that sometimes we intuitively feel that one sentence has a stronger meaning (or is more informative) than another, but the former does not directly entail the latter.

In (21a), *Miranda is 1.65* m *tall* means that the height of Miranda reaches the measurement of 1.65 m, i.e., $\lambda w.\textsc{height}(\textsc{Miranda})(w) \geq 1.65$ m. Thus it does entail *Miranda is 1.60* m *tall* – $\lambda w.\textsc{height}(\textsc{Miranda})(w) \geq 1.60$ m.

The two sentences mentioned in footnote 1 (repeated here as (22)) should be interpreted in a way parallel to the two sentences in (21a). Actually we do have a natural intuition that (22a) has a stronger meaning than (22b). However, it is evident that (22a) does not directly entail (22b).

(22) a. John is above 6 feet tall. $\lambda w.\textsc{height}(\textsc{John})(w) \subseteq [6', +\infty)$
b. John is between 4 and 5 feet tall. $\lambda w.\textsc{height}(\textsc{John})(w) \subseteq [4', 5']$

Under the current degree-QUD-based view on maximality of informativeness, I tease apart (i) the height measurement (typically with units like feet, meters, etc.) and (ii) the degree of tallness. Presumably, items of different comparison class can share the same scale for height measurement (e.g., the height of humans, giraffes, and mountains can be measured along the same scale and with the same units). However, the degrees of tallness and the comparison between them hinge on the notion of comparison class (e.g., toddlers are usually compared with other toddlers in terms of tallness). Thus it is evident that although '$\lambda w.\text{HEIGHT}(\text{JOHN})(w) \subseteq [6', +\infty)$' does not entail '$\lambda w.\text{HEIGHT}(\text{JOHN})(w) \subseteq [4', 5']$', under a degree QUD like *to what extent is John tall*, the measurement '$[6', +\infty)$' represents a higher degree in addressing this degree QUD and is thus more informative than '$[4', 5']$'. Therefore, our intuition that (22a) has a stronger meaning than (22b) can be accounted for.

For the core example (1), it is also worth noting that under the scenario of Fig. 1, although *exactly 3 boys saw exactly 5 movies* holds true, *exactly 1 boy saw exactly 4 movies* does not hold true (in Fig. 1, no boys saw more than 3 movies). Thus, this example also shows that it is problematic to build informativeness ordering directly upon the entailment relation between uttered sentences and their alternatives (derived by replacing uttered numbers with other numbers). However, *exactly 3 boys saw exactly 5 movies* does indicate a higher film consumption level (or a more prosperous situation) than the consumption level indicated by *exactly 1 boy saw exactly 4 movies*. Thus, the uttered sentence indicates a higher informativeness in addressing an underlying QUD than its alternatives. In this sense, with the use of a degree QUD, the current proposal provides a more generalized view on informativeness than an entailment-based one.

6 Extension: QUD-Based Informativeness and *even*

Beyond cumulative-reading sentences (and measurement sentences like (22)), here I use the case of *even* to show a broader empirical coverage of the proposed QUD-based view on maximality of informativeness (see also Zhang 2022).[3]

According to the traditional view on *even*, its use brings two presuppositions (and presuppositions are considered a kind of entailment): (i) **entity-based additivity** (see (23a)) and (ii) **likelihood-based scalarity** (see (23b)).

(23) (It's not the case that) even $[\text{Mary}]_F$ came.

a. (23) $\models$ Someone other than Mary came.

b. (23) $\models$ Compared to others, Mary was unlikely to come.

However, it seems that the notion of entailment is too strong to characterize the meaning inferences with regard to the use of *even*. As illustrated by an example from Szabolcsi (2017), under the given scenario in (24), the use of an *even*-sentence in (25) is perfectly natural, but it challenges the traditional entailment-based view on our natural inferences for *even*-sentences. First,

[3] I thank an anonymous reviewer for raising this issue, which has led me to see this kind of connection that I missed before.

as shown in (25a), the presuppositional requirement of additivity is not met, because Eeyore was the only one who took a bite of thistles and spit them out. Second, if no one other than Eeyore took a bite of thistles, it seems also questionable to claim that the likelihood of the truth of the prejacent is lower than that of *X spit thistles out* ($X \in \text{Alt}(\text{Eeyore})$), as shown in (25b).

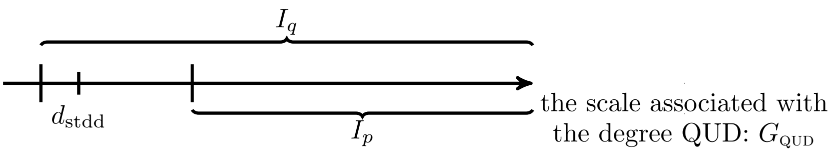
(24) **Scenario**: Imagine Pooh and friends coming upon a bush of thistles. Eeyore (known to favor thistles) takes a bite but spits it out.

(25) (Those thistles must be really prickly!) Even [Eeyore]$_F$ spit them out!
  a. (25) $\not\models$ Someone other than Eeyore spit thistles out.
  b. (25) $\not\models$ Compared to others, Eeyore was unlikely to spit thistles out.

Zhang (2022) proposes a degree-QUD-based analysis for the use of *even* (see Greenberg 2018 for a similar view). The use of *even* is always based on a contextually salient degree QUD (for (25), *how prickly are those thistles*). The prejacent of *even* (here *Eeyore spit those thistles out*) provides information to resolve this degree QUD with an increasingly positive answer, and compared with alternatives, this prejacent is also considered maximally informative in resolving this degree QUD. I.e., here compared to *X spit those thistles out* ($X \in \text{Alt}(\text{Eeyore})$), *Eeyore spit those thistles out* is maximally informative in resolving the degree question *how prickly are those thistles.*

Therefore, as illustrated in (26), the presupposition of *even* contains two parts. First, in all the accessible worlds where the prejacent is true, the range of the prickliness measurement of thistles, I_p, exceeds the threshold d_{stdd} (i.e., the degree QUD is resolved by the prejacent with a positive answer). Second, compared to I_q (i.e., the range of the prickleness measurement of thistles informed by an alternative statement *X spit thistles out*), I_p is maximally informative.[4]

(26) The degree-QUD-based presupposition of *even* proposed by Zhang (2022):



$$I_p = \text{Max}_{\text{info}}[\lambda I.[\forall w' \in \text{Acc}(w) \cap p[G_{\text{QUD}}(x_{\text{QUD}})(w') \subseteq I]]],$$
$$I_q = \text{Max}_{\text{info}}[\lambda I.[\forall w'' \in \text{Acc}(w) \cap q[G_{\text{QUD}}(x_{\text{QUD}})(w'') \subseteq I]]].$$

It is interesting to see that our interpretation for both cumulative-reading sentences and focus-related sentences can be based on the same degree-QUD-based mechanism of informativeness and demonstrate the maximality of informativeness.

[4] In Zhang (2022), I implement my analysis based on intervals, instead of degrees (see also Abrusán 2014; Schwarzschild and Wilkinson 2002; Zhang and Ling 2021).

7 Conclusion

Starting from the discussion on our intuitive interpretation for two kinds of cumulative-reading sentences, this paper proposes a degree-QUD-based view on the maximality of informativeness. The informativeness of a sentence basically stands for how the sentence resolves a contextually salient degree QUD.

For cumulative-reading sentences like *Exactly three boys saw exactly five movies*, its informativeness means the degree information in addressing *how high the film consumption level is among boys*, and the uttered numbers reflects mereological maximality. Then for cumulative-reading sentences like *In Guatemala, at most 3% of the population own at least 70% of the land*, its informativeness means rather the degree information in addressing *how skewed wealth distribution is in Guatemala*, and the uttered numbers reflects the maximality of the ratio between land and their owner population.

It seems that the current QUD-based view on the maximality of informativeness can overcome some issues that challenge the existing entailment-based view on informativeness and provide a broader empirical coverage. A further development of the current proposal to account for other related phenomena, especially with regard to the interpretation of numerals and focus items (e.g., *even*), as well as a more detailed discussion on its theoretical implications are left for future research.

References

Abrusán, M.: Weak island effects. OUP (2014). https://doi.org/10.1093/acprof:oso/9780199639380.001.0001

Brasoveanu, A.: Modified numerals as post-suppositions. J. Semant. **30**, 155–209 (2013). https://doi.org/10.1093/jos/ffs003

Bumford, D.: Split-scope definites: relative superlatives and Haddock descriptions. Linguist. Philos. **40**, 549–593 (2017). https://doi.org/10.1007/s10988-017-9210-2

Charlow, S.: Post-suppositions and semantic theory. J. Semant. (2017, in Press). https://ling.auf.net/lingbuzz/003243

von Fintel, K., Fox, D., Iatridou, S.: Definiteness as maximal informativeness. In: Crnič, L., Sauerland, U. (eds.) The Art and Craft of Semantics: A Festschrift for Irene Heim, vol. 2, pp. 175–179. Cambridge (2014). https://semanticsarchive.net/Archive/jZiNmM4N/

Greenberg, Y.: A revised, gradability-based semantics for even. Nat. Lang. Semant. **26**(1), 51–83 (2018) https://doi.org/10.1007/s11050-017-9140-0

Grindrod, J., Borg, E.: Questions under discussion and the semantics/pragmatics divide. Philos. Q. **69**(275), 418–426 (2019). https://doi.org/10.1093/pq/pqy058

Hendriks, P., De Hoop, H.: Optimality theoretic semantics. Linguist. Philos. **24**, 1–32 (2001). https://www.jstor.org/stable/25001802

Krifka, M.: At least some determiners aren't determiners. In: Tuner, K. (eds) The semantics/pragmatics interface from different points of view, vol. 1, 257–291. Elsevier (1999). https://semantics.uchicago.edu/kennedy/classes/w14/implicature/readings/krifka99.pdf

Roberts, C.: Information structure: towards an integrated formal theory of pragmatics. Semant. Pragmat. **5**, 6:1–69 (2012). https://doi.org/10.3765/sp.5.6

Schlenker, P.: Informativity-based maximality conditions. Snippets. **26**, 18–19 (2012). https://www.ledonline.it/snippets/allegati/snippets26007.pdf

Schwarzschild, R., Wilkinson, K.: Quantifiers in comparatives: a semantics of degree based on intervals. Nat. Lang. Semant. **10**, 1–41 (2002). https://doi.org/10.1023/A:1015545424775

von Stechow, A.: Comparing semantic theories of comparison. J. Semant. **3**, 1–79 (1984). https://doi.org/10.1093/jos/3.1-2.1

de Swart, H.: Indefinites between predication and reference. Proc. SALT **9**, 273–297 (2006). https://doi.org/10.3765/salt.v9i0.2823

Szabolcsi, A.: Strategies for scope taking. In: Szabolcsi, A. (eds) Ways of Scope Taking, pp. 109–154. Springer, Cham (1997). https://doi.org/10.1007/978-94-011-5814-5_4

Szabolcsi, A.: Additive presuppositions are derived through activating focus alternatives. Proc. Amsterdam Colloquium **21**, 455–464 (2017). https://semanticsarchive.net/Archive/jZiM2FhZ/AC2017-Proceedings.pdf

Umbach, C.: Why do modified numerals resist a referential interpretation. Proc. SALT **15**, 258–275 (2006). https://doi.org/10.3765/salt.v15i0.2931

Zhang, L.: The presupposition of even. Proc. SALT **32**, 249–269 (2022). https://doi.org/10.3765/salt.v1i0.5355

Zhang, L.: Post-suppositions and uninterpretable questions. Proc. TLLM, 167–187 (2023). https://doi.org/10.1007/978-3-031-25894-7_9

Zhang, L., Ling, J.: The semantics of comparatives: A difference-based approach. J. Semant. **38**, 249–303 (2021). https://doi.org/10.1093/jos/ffab003