The semantics and pragmatics of multi-head comparatives

Linmin Zhang, zhanglinmin@gmail.com, the version of October 13, 2024

Abstract

A **multi-head comparative**, e.g., *Fewer people own more wealth in this country than anywhere else* (see also Chomsky 1981/1993, von Stechow 1984), typically contains two comparative expressions and only one *than*-clause.

I argue that there is a connection between a **cumulative-reading sentence** like *at most 3% of the population own at least 70% of the wealth* (see also Krifka 1999, Brasoveanu 2013) and the above multi-head comparative. This connection is parallel to that between a measurement sentence (e.g., *Mary is exactly 6 feet tall*) and a regular comparative (e.g., *Sue is taller than Mary is*): the former means a **measurement**, while the latter means **comparison** between measurements. Thus, semantically, a multi-head comparative is the comparative form of a multi-measurement sentence, expressing comparisons along multiple dimensions simultaneously.

Pragmatically, I show that the above multi-head comparative (along with its corresponding cumulative-reading sentence) addresses a single issue: how skewed wealth distribution is. Thus in a multi-head comparative, comparisons along multiple dimensions are not interpreted independently. There is an interplay between changes along two dimensions, which resolves a single context-dependent issue. I propose the notion of 'degree-QUD-based informativeness' to analyze this interplay.

Keywords: Comparatives, Multi-head comparatives, Multi-measurement, Cumulative reading, Degree questions, Multi-degree questions, QUD, Dynamic semantics, Maximality, Informativeness, Comparative correlatives

1 Introduction

A **multi-head comparative** typically contains two comparative expressions and only one *than*-clause, as illustrated by the examples (1)–(4), all from von Stechow (1984).

(1) More silly lectures have been given by more boring professors – than I would have expected. (p.42: (136a); originally from Chomsky 1981/1993: p.81, §2.4.4, (6ii)

(2) More dogs ate more rats than cats ate mice. (p.43: (140))

(3) Less land produces more corn than ever before. (p.46: (156))

(4) No airline saves you more money in more ways than Delta. (p.46: (157))

Although multi-head comparatives have been noticed for decades in formal syntax and semantics (see e.g., Chomsky 1981/1993, von Stechow 1984), a thorough investigation is still missing. On the one hand, von Stechow (1984) suggests that the semantics of multi-head comparatives should be a natural and ready extension of our existing understanding of comparatives:

> 'I don't think that these cases represent genuine semantic problems. They can essentially be treated with the methods we already have at our disposal. But it took me quite a while to realize this, because the examples are conceptually rather complicated and are neglected in the literature.'
>
> (von Stechow 1984, pp.41–42)

On the other hand, it has been questioned whether multi-head comparatives are, after all, interpretable and sensible (see e.g., Hendriks 1994, Hendriks and De Hoop 2001). Even if von Stechow (1984) denies that they present a real challenge, in his quotation above, he still characterizes them as 'conceptually rather complicated'.

Corpus search from *the Corpus of Contemporary American English* (https://www.english-corpora.org/coca/, Davies 2008-) shows that there are indeed naturally occurring sentences that contain multiple comparative expressions:

(5) Henry Ford's assembly line in 1913 increased productivity. Fewer people made more stuff. (2002, NEWS: Atlanta Journal Constitution)

(6) Fewer and fewer people own more and more of the country's land. (2001, MAG: Christian Century)

(7) In the four years of the (Marshall) Plan, the Marshall agency spent \$13.5 billion in 16 countries. Fewer people spent more money in that agency than ever before. (1998, TV: Cold War)

(8) As I moved through the years, I saw that I would make more money in more skilled positions. (2012, BLOG: pjmedia.com)

However, for a sentence, containing multiple comparative expressions does not guarantee its being a multi-head comparative, as evidenced by examples like (9a). Although (9a) contains two comparative expressions while (9b) only one, we intuitively feel that (9a) and (9b) are interpreted in a similar way. Both mainly involve a comparison of school numbers, not a comparison of the length of school year.

(9) a. More schools choose a longer school year than before.
(adapted from https://www.aps.edu/news/archives/news-from-2021-2022/9-more-aps-schools-choose-a-longer-school-year)
b. More schools choose a long school year than before.

Empirically, what characterizes genuine multi-head comparatives? Theoretically, how to provide a formal analysis? The current paper aims to answer these two questions.

In a nutshell, I propose that **semantically**, a genuine multi-head comparative is the comparative form of a corresponding sentence that expresses **measurements along multiple dimensions simultaneously** (e.g., a cumulative-reading sentence).

I propose that **pragmatically**, this kind of **simultaneous** comparisons / measurements along multiple dimensions address a **single** context-dependent degree QUD (Question-under-discussion, see Roberts 1996/2012). For example, (1) essentially addresses an underlying degree QUD like how bad education quality is.

I develop the notion of '**degree-QUD-based informativeness**' to analyze the simultaneity of multi-dimensional measurement / comparison. Specifically, I propose to tease apart (i) **context-independent measurements**, which can often be expressed as numerical values, and (ii) **context-dependent QUD-based degrees**, which address whether the increase/decrease of measurements leads to higher informativeness.

In (10), measurements and comparisons provide context-independent height information to resolve questions like *what is the height of Mary/Sue/Jane/Lily*. Height information does not directly determine informativeness: e.g., '$> 6'$' does not entail and is thus not considered more informative than '$= 6'$'.

However, in addressing an underlying concern that correlates with the degree of tallness reached, if someone reaches the tallness degree associated with the measurement '$> 6''$', it is guaranteed that they reach the tallness degree associated with '$= 6''$'. Thus the increase of height values leads to the increase of informativeness in addressing this underlying degree QUD (see (10a)). Conversely, in addressing the degree of shortness reached, it is the decrease of height values that leads to the increase of informativeness (see (10b)).

In comparatives, the measurement of the comparison target at the matrix level (e.g., *Sue* in (10a) and *Lily* in (10b)) is associated with higher informativeness in addressing an underlying degree QUD than the measurement provided by the *than*-clause.

(10) a. **QUD on the degree of tallness reached:** height $\uparrow \rightsquigarrow$ informativeness $\uparrow$

$$\underbrace{\text{Mary is exactly 6 feet tall}}_{\text{the height of Mary}} \rightsquigarrow \text{Sue is } \underbrace{\text{taller } \underbrace{\text{than Mary is } \cancel{d\text{-tall}}}_{\text{the height of Mary}}}_{\text{the height of Sue} \rightsquigarrow \text{higher informativeness}}$$

b. **QUD on the degree of shortness reached:** height $\downarrow \rightsquigarrow$ informativeness $\uparrow$

$$\underbrace{\text{Jane is exactly 5 feet tall}}_{\text{the height of Jane}} \rightsquigarrow \text{Lily is } \underbrace{\text{shorter } \underbrace{\text{than Jane is } \cancel{d\text{-short}}}_{\text{the height of Jane}}}_{\text{the height of Lily} \rightsquigarrow \text{higher informativeness}}$$

Similarly, in a multi-head comparative, measurements and comparisons along multiple dimensions provide context-independent measurement information to resolve multi-*how-many/much* questions like *how many cats ate how many mice* (see (11)) or *how many people own how much wealth* (see (12)). However, it is not multi-dimensional measurement information, but rather an underlying degree QUD that determines informativeness. For (11), the underlying degree QUD (e.g., *as predators, how successful are dogs / cats*) determines that parallel increases along both dimensions lead to higher informativeness, while for (12), the underlying degree QUD (e.g., *how wealth distribution is skewed*) determines that the decrease of the owner population along with the increase of wealth quantity leads to higher informativeness. Then just like in (10), in multi-head comparatives, the measurements at the matrix level resolve the underlying degree QUD with higher informativeness than the measurements provided by the *than*-clause.

(11) **Underlying degree QUD**: $\underline{\text{\#predators} \uparrow, \text{\#prey} \uparrow} \rightsquigarrow$ informativeness $\uparrow$

$$\underbrace{\text{3 cats ate 5 mice}}_{\text{How many cats ate how many mice}} \rightsquigarrow$$

$\underbrace{\text{More dogs ate more rats } \underbrace{\text{than } \cancel{m\text{-many}} \text{ cats ate } \cancel{n\text{-many}} \text{ mice}}_{\text{How many cats ate how many mice}}}_{\text{How many dogs ate how many rats} \rightsquigarrow \text{ higher informativeness}}$

(12) **Underlying degree QUD**: $\underline{\text{\#owners} \downarrow, \text{\#wealth} \uparrow}$ $\rightsquigarrow$ informativeness $\uparrow$

$\underbrace{\text{3\% of the population own 70\% of the wealth elsewhere}}_{\text{How many people own how much wealth elsewhere}} \rightsquigarrow$

$\underbrace{\text{Fewer people own more wealth here } \underbrace{\text{than } \cancel{m\text{\% of the population own } n\text{\% of the wealth}} \text{ elsewhere}}_{\text{How many people own how much wealth elsewhere}}}_{\text{How many many people own how much wealth here} \rightsquigarrow \text{ higher informativeness}}$

Thus, multi-head comparatives are a combination of multi-measurement sentences (e.g., cumulative-reading sentences) and comparatives, using an interplay between multi-dimensional comparisons to address an underlying degree QUD.

The proposed analysis supports von Stechow (1984)'s view that multi-head comparatives are not a challenge to our existing understanding of comparatives and 'can be treated with the methods we already have at our disposal'. This analysis is achieved by (i) connecting multi-head comparatives with well-studied phenomena (i.e., comparatives and cumulative-reading sentences) and (ii) teasing apart context-independent measurements and context-dependent degree QUDs.

Eventually, multi-head comparatives (along with multi-measurement sentences) demonstrate how natural language supports the expression of complex mathematical operations that range from one-dimensional to multi-dimensional comparison and how the interpretation involved is guided by an underlying goal and integrates multi-dimensional information into a holistic whole.

The rest of the paper is organized as follows. Based on von Stechow (1984), Hendriks (1994), Marques (2005), and Oda (2008a,b), §2 presents empirical observations. §3 connects multi-head comparatives with cumulative-reading sentences. I combine the studies of Krifka (1999), Brasoveanu (2013), Zhang and Ling (2021), and Zhang (2023) to propose the key ideas of the current analysis (see the sketches in (11) and (12)). A formal implementation that combines interval semantics and dynamic semantics is presented in §4. §5 addresses theoretical implications with regard to the notions of comparison and informativeness, and includes a brief discussion on comparative correlatives. §6 concludes.

2 Empirical observations

Empirical observations presented in this section will start with von Stechow (1984)'s discussion on how to interpret typical multi-head comparatives (§2.1). Then I discuss cases deviating from typical multi-head comparatives and whether / to what extent they are interpretable (§2.2). The discussion in this section aims to show the empirical scope of this paper and pave the way for the proposed analysis. Observations related to the connection between multi-head comparatives and cumulative-reading sentences will be further discussed in §3.

2.1 Typical multi-head comparatives and their interpretation

For a typical multi-head comparative like (2), von Stechow (1984) describes what it intuitively means and what it does not mean.

As shown in (13), intuitively, the most readily available reading of sentence (2) means **two comparisons**, each along a distinct dimension (see von Stechow 1984, p.43: (141)). Along the dimension of **agent cardinalities**, dogs that ate rats are compared with cats that ate mice (see (13a)). Along the dimension of **theme cardinalities**, rats eaten by dogs are compared with mice eaten by cats (see (13b)).

(13) **von Stechow (1984): what a multi-head comparative intuitively means:**

$\underbrace{\text{More dogs}}_{\text{agent}}$ ate $\underbrace{\text{more rats}}_{\text{theme}}$ than $\underbrace{\text{cats}}_{\text{agent}}$ ate $\underbrace{\text{mice}}_{\text{theme}}$. (= (2))

a. The comparison of **agent cardinalities**: |dogs| vs. |cats|
The number of dogs that ate rats > the number of cats that ate mice

b. The comparison of **theme cardinalities**: |rats| vs. |mice|
The number of rats eaten by dogs > the number of mice eaten by cats

For this reading sketched in (13), multiple comparisons are along distinct dimensions, and these dimensions (and comparisons along each dimension) are mutually independent (see (13a) vs. (13b)).

Furthermore, von Stechow (1984) notes that there is a **mutual restriction** between the agent and the theme, at the matrix-clause level and the *than*-clause level respectively.

In interpreting (2)/(13), we do not count all the dogs and cats in our context and then compare the cardinalities of the absolutely maximal dog-sum and cat-sum. Rather, we count all the dogs that ate rats and all the cats that ate mice and then compare the cardinalities of the restricted dog-sum and cat-sum. Similarly, it is also for the restricted rat-sum (those eaten by dogs) and mouse-sum (those eaten by cats) that their cardinalities undergo comparison. This mutual restriction is also the hallmark of cumulative reading and will be discussed in §3.

According to von Stechow (1984), (14) is not a good characterization of our intuitive interpretation of (2) (see von Stechow 1984, p.43: (142)). In contrast to (13), which expresses two mutually independent comparisons along two dimensions, (14) only expresses **one comparison**: the comparison between the cardinality of all dog-rat pairs involved in eating-events and that of all cat-mouse pairs involved in eating events.

(14) **von Stechow (1984): what a multi-head comparative like (2) does not mean:**
The number of pairs $\langle x, y\rangle$ > the number of pairs $\langle x', y'\rangle$
(where $\textsc{dog}(x), \textsc{rat}(y), \textsc{cat}(x'), \textsc{mouse}(y'), \textsc{eat}(x, y), \textsc{eat}(x', y')$)

This event-counting-based comparison reading in (14) is too weak, which can be confirmed with a scenario like (15). Under the scenario (15), we intuitively judge sentence (2) false. However, the reading shown in (14) is actually true under this scenario, which is at odds with our intuition.

(15) Scenario: here are all the animal pairs involved in eating events:
$\langle \text{Dog}_1, \text{Rat}_1\rangle$, $\langle \text{Dog}_1, \text{Rat}_2\rangle$, $\langle \text{Dog}_1, \text{Rat}_3\rangle$, $\langle \text{Cat}_1, \text{Mouse}_1\rangle$, $\langle \text{Cat}_2, \text{Mouse}_2\rangle$
Overall, 1 dog ate 3 rats, and 2 cats ate 2 mice between them. Thus here
$|\langle x, y\rangle| = 3, |\langle x', y'\rangle| = 2, \therefore, |\langle x, y\rangle| > |\langle x', y'\rangle|$.

2.2 Deviations from typical multi-head comparatives like (2)

Hendriks (1994) (see also Hendriks and De Hoop 2001) points out that von Stechow (1984)'s view on multi-head comparatives would lead to over-generation: Not all sentences containing multiple comparative expressions make sense in the way sketched in (13) (i.e., expressing multiple comparisons along distinct dimensions), and some sentences with a similar pattern to (2)/(13) are in fact intuitively nonsensical.

Here I discuss 6 types of deviations from (2)/(13), investigating whether / to what extent they make sense and what sense they make.

2.2.1 Comparative expressions with opposite directions of inequalities

von Stechow (1984)'s analysis of (2)/(13) is expected to be extendable to sentences like (16). However, according to Hendriks (1994), we intuitively feel that (16) is nonsensical.

(16) *$\underbrace{\text{Fewer dogs}}_{\text{agent}}$ ate $\underbrace{\text{more rats}}_{\text{theme}}$ than $\underbrace{\text{cats}}_{\text{agent}}$ ate $\underbrace{\text{mice}}_{\text{theme}}$. (Hendriks 1994: (6))

The analysis à la von Stechow (1984): two comparisons (Hendriks 1994: (6′))

a. Along the dimension of agent cardinalities:
The number of dogs that ate rats < the number of cats that ate mice

b. Along the dimension of theme cardinalities:
the number of rats eaten by dogs > the number of mice eaten by cats.

Is the degradedness of (16) due the opposite directions of the two inequalities (see (16a) vs. (16b))? The existence of data like (3) (and naturally occurring examples like (6) and (7)) indicates that the answer is *no*. (3) (repeated in (17)) is similar to (16) in involving two comparisons of opposite directions, but (3)/(17) is an intuitively natural and acceptable multi-head comparative.

(17) <u>Less</u> land produces <u>more</u> corn <u>than</u> ever before. (= (3))

Then why does our intuition for (16) contradict the prediction of von Stechow (1984)? Hendriks (1994) argues that the two comparisons involved in the interpretation of a multi-head comparative cannot be mutually independent and proposes the analysis in (18) as the meaning of the degraded sentence (16).

(18) **Hendriks (1994)'s analysis of (16)** (see Hendriks 1994: (8)):
the number of dogs x such that '|rats eaten by x| > |mice eaten by y|' <
the number of cats y such that '|rats eaten by x| > |mice eaten by y|'

In (18), comparison along one dimension is embedded within comparison along another dimension. The embedded comparison '|rats eaten by x| > |mice eaten by y|' restricts the items x and y, and x and y are further compared along the dimension of cardinality. Hendriks (1994) claims that this restriction '|rats eaten by x| > |mice eaten by y|' makes x and y mutually dependent and results in infinite regress: the definition of x relies on the definition of y, and vice versa. Thus this mutual dependency cannot be derived compositionally, and the meaning of (16) is nonsensical and uninterpretable.[1]

Hendriks (1994)'s analysis of the degradedness of (16) in (18) is dubious in that sentences like (2)/(13) or (3)/(17) are intuitively natural and interpretable, indicating that their semantic derivation does not suffer from the potential infinite regress problem.

Thus multi-head comparatives are not always incompatible with comparisons of opposite directions. The degradedness of examples like (16) needs a different explanation.

Meier (2001) makes such a comment: 'The fact that (17a) (= (16) in the current paper) is somehow odd is just a reflex of the fact that it is hard to imagine a situation in which the comparison construction might be relevant (p.356).' Indeed, (19) illustrates that a sentence similar to (16) sounds natural when uttered in an appropriate context.

(19) Context: A zoo keeper is talking about the cost of feeding animals. A lion eat between 5 to 7 kg of meat a day, while a chimpanzee typically eats fruit, and the amount is between 1 to 4 kg a day. Feeding and keeping lions is more costly.[2]
Fewer lions eat more meat than chimpanzees eat fruit. (cf. (16))

Under the context in (19), our interpretation of this multi-head comparative can be naturally accounted for by an analysis à la von Stechow (1984). (19) means two comparisons, one along the dimension of animal cardinalities and the other along the amount of consumed food, and opposite directions of inequalities indicate how feeding lions has a lower value than feeding chimpanzees along a scale of cost-effetiveness.

2.2.2 Multi-head comparatives with proportional reading or distributive reading

In the above discussion, von Stechow (1984)'s analysis shows that multi-head comparatives involve two comparisons that are along two dimensions: one measures and compares the agents, and the other measures and compares the themes. Moreover, in the above examples, measurements typically count the **cardinality or total amount** of the agents / themes. But multi-head comparatives are not necessarily based on the measurements / comparisons of cardinalities or total amounts, as evidenced by the existence of **proportional-reading** and **distributive-reading** multi-head comparatives.

[1]In §3 and §4, I will show that this kind of mutual restriction does not challenge semantic compositionality. Works like Brasoveanu (2013) have provided a solution (see also Charlow 2017). In addition to the literature on cumulative reading, Bumford (2017) investigates the mutual definition of uniqueness in Haddock descriptions (e.g., *the rabbit in the hat*, see Haddock 1987) and provides an analysis.

[2]I thank Takeo Kurafuji for discussing the example (16) with me and providing this scenario in (19).

(20) and (21) are naturally occurring examples of multi-head comparatives with a proportional reading. According to our world knowledge, the measurements involved in interpreting (20) and (21) are most naturally based on percentages, instead of cardinalities. In this sense, multi-head comparatives are like regular comparatives in being flexible with various kinds of measurement units (see (22)).

(20) The trend is indisputable: <u>Fewer</u> people own <u>more</u> of the overall wealth, and <u>fewer</u> companies own <u>more</u> market share.
(https://www.deseret.com/opinion/2020/9/14/21436415/guest-opinion-america-capitalism-strengths-dark-side-too-far-inequality-divisiveness-wealth-gap)

(21) <u>Fewer</u> people own <u>more</u> of the land in Brazil <u>than</u> anywhere else in the world.
(https://glenmorangie.newint.org/features/2003/01/05/cutting)

(22) <u>More</u> people are hired here <u>than</u> elsewhere. **cardinalities ✓; percentages ✓**

Flexibility of measurement units can also be observed from the phenomenon of distributive-reading multi-head comparatives. Marques (2005) and Oda (2008a,b) have mentioned this kind of phenomenon in Portuguese and Japanese.

In (23) and (24), the measurement and comparison along one dimension is based on cardinalities (i.e., number of predators in (23) and number of companies in (24)), while the measurement and comparison along the other dimension is based on per-unit values (i.e., number of preys per predator in (23) and amount of corporate tax per company in (24)). The distributive meaning is overtly expressed by the use of *sorezore* in (23) but covertly inferred from our world knowledge in interpreting (24).

(23) **Distributive-reading multi-head comparative in Japanese**

San-biki-no	neko-ga	**sorezore**	yon-hiki-no	hatukanezumi-o	tabeta	<u>yorimo</u>
3-CL-GEN	cat-NOM	each	4-CL-GEN	mouse-ACC	ate	THAN

(<u>motto</u>)	<u>takusanno</u>	inu-ga	**sorezore**	(<u>motto</u>)	<u>takusanno</u>	dobunezumi-o
(more)	many	dog-NOM	each	(more)	many	rat-ACC

tabeta.
ate

Lit. 'More dogs ate more rats each than three cats ate four mice each.'
⇝ There are 3 cats and each of them ate 4 mice. There are more than 3 dogs and each of them ate more than 4 rats. (Oda 2008b: (62) and (63))

(24) **Distributive-reading multi-head comparative in Portuguese**

Eles querem que mais empresas paguem menos IRC
they want that more companies pay less corporate-tax

'They want more enterpreises to pay less corporate tax.' (Marques 2005: (65))

There are also naturally occurring distributive-reading multi-head comparatives in English, as illustrated in (25). Based on our world knowledge, *more of their time* in (25) should be measured and compared per person, resulting in a distributive reading.

(25) The world would be a more positive place to be if more people spent more of their time focusing on what they're thankful for.
(https://schizanthusnerd.com/2018/04/01/everyday-gratitude/)

2.2.3 Comparative expressions beyond plural/mass nouns

von Stechow (1984) suggests that comparative expressions in a multi-head comparative are restricted to plural or mass nouns (e.g., *more dogs* and *more rats* in (2), *less land* in (3)). A sentence like (26) is ungrammatical, because here comparative expressions *greater* and *better* are not plural or mass nouns. Hendriks (1994) also presents an ungrammatical case like (27), in which the comparative expression *higher* is not a plural or mass noun.

(26) *A greater man would be a better man than Otto. (von Stechow 1984: p.46, (158))

(27) *$\underbrace{\text{More doors}}_{\text{items}}$ are $\underbrace{\text{higher}}_{\text{linear size}}$ than $\underbrace{\text{windows}}_{\text{items}}$ are $\underbrace{\text{wide}}_{\text{linear size}}$. (Hendriks 1994: (5))

According to von Stechow (1984), in a multi-head comparative, each comparative expression needs to be associated with an overt or silently reconstructed *than*-clause. For a sentence like (26), von Stechow (1984) claims that the reconstruction gives rise to two possibilities in interpreting the *than*-phrase (see (28a) and (28b)), and 'the low acceptability of (26) shows that our grammar doesn't work that way. It seems, then, that the restrictions for the reconstruction of a *than*-phrase are rather syntactic than semantic.'

(28) *A greater man (**than Otto**) would be a better man **than Otto**.
   a. than Otto is a great man
   b. than Otto is a good man

von Stechow (1984)'s reconstruction-based explanation in (28) is challenged by examples like (29), which contains two comparative expressions that are not plural or mass nouns. The grammaticality of (29) shows that there is in principle no problem in associating two comparative expressions with the same *than*-clause/phrase. The interpretation of (29) involves two comparisons along the dimensions of greatness and goodness, and the recovery of the elided gradable adjectives in the *than*-clause works like in a regular comparative that involves only one comparative expression.

(29) He is a greater and better man than I am. ⇝ than I am ~~great and good~~

Furthermore, von Stechow (1984)'s empirical generalization, i.e., the 'plural/mass nouns' restriction, is also challenged by the existence of examples like (30)–(32). In these multi-head comparatives, at least one of the comparative expressions is not a plural or mass noun: *faster* in (30), *more varied* and *better* in (31), and *better* in (32).

(30) Nowadays, more goods are carried faster (than before).
(Hendriks and De Hoop 2001: p.10, (13))

(31) Aldi's pork selection was a little more varied with better value than what I saw here. (https://www.insider.com/aldi-vs-lidl-review-differences-which-better-photos-2021-5)

(32) Hydrogen-powered cars also provide slightly more driving range with better energy density than batteries and fuel. (https://www.whichcar.com.au/car-advice/hydrogen-cars-v-electric-cars-australia)

I will further discuss unacceptable examples (26) and (27) in §2.2.4 and §2.2.6.

2.2.4 The lack of parallelism between the matrix and the *than*-clause

von Stechow (1984) points out that (33) is intuitively unacceptable as a multi-head comparative. According to von Stechow (1984), the analysis of a multi-head comparative requires the silent reconstruction of a *than*-clause, but (33) is different from (1) (repeated here as (34)) in that reconstructing the *than*-clause for the first comparative expression in (33) is impossible (see (35)).

(33) *More silly lectures have been given by more boring professors than I met yesterday. (Chomsky 1981/1993: p.81, §2.4.4, (6iii))

(34) More silly lectures have been given by more boring professors – than I would have expected. (=(1))
⇝ More silly lectures (**than I would have expected that silly lectures would be given by boring professors**) have been given by more boring professors **than I would have expected (that silly lectures would be given by boring professors)**.

(35) *More silly lectures (**than I met yesterday**) have been given by more boring professors **than I met yesterday**. Reconstructing the *than*-clause for (33)

von Stechow (1984)'s account for the ungrammaticality of (33) is based on the ungrammaticality of (35): *meet* is compatible with nouns like *professors*, but not nouns like *lectures*. If *meet* in (33)/(35) is replaced by *evaluate*, a verb compatible with both *lectures* and *professors*, then a sentence like (37), which is parallel to (35), is acceptable.

If a multi-head comparative indeed involves reconstructing a *than*-clause for the first comparative expression, we would predict that (36) should be as acceptable as (37). But this prediction is not borne out. Intuitive, (36) is as weird as (33).

(36) *More silly lectures have been given by more boring professors than I evaluated last year.

(37) Context: As the academic committee chair, I evaluated 5 courses last year, and I also evaluated 6 professors. To my surprise, the teaching quality is now dropping dramatically, in terms of both the number of silly lectures and the number of boring professors.
More silly lectures than I evaluated last year have been given by more boring professors than I evaluated last year.

Thus the contrast between the ungrammatical multi-head comparative (36) and the acceptable sentence in (37) argues against von Stechow (1984)'s reconstruction-based analysis. The interpretation of multiple comparisons in a multi-head comparative should not be based on a distinct, independent *than*-clause for each comparative expression.

Rather, the comparison standards needed in the interpretation of multiple comparisons should all be recoverable from the same *than*-clause. In other words, parallelism is needed between the matrix and the *than*-clause in providing information on multiple measurements for multi-dimensional comparison.

As illustrated in (38), for a regular, one-dimensional comparative (see (38a)), both the matrix and the *than*-clause provide single-dimensional measurements. Similarly, for a multi-head comparative that involves multi-dimensional comparisons (see (38b)), both the matrix and the *than*-clause provide parallel multi-dimensional measurements.

(38) **Parallelism between what undergoes comparisons**:

a. $\underbrace{\text{Sue is taller } \underbrace{\text{than Mary is } \cancel{\text{tall}}}_{\text{comparison standard: Mary's height}}}_{\text{comparison target: Sue's height}}$ **One-dimensional**

b. $\underbrace{\text{More dogs ate more rats } \underbrace{\text{than cats ate mice}}_{\text{standards: how many cats ate how many mice}}}_{\text{targets: how many dogs ate how many rats}}$ **Two-dimensional**

Thus sentences like (33) and (36) are weird because their *than*-clause (i.e., *than I met ~~d-many professors~~ yesterday* in (33), *than I evaluated ~~d-many professors~~ last year* in (36)) does not provide multi-dimensional measurements, failing to support the parallelism between the matrix and the *than*-clause to conduct multi-dimensional comparison.

In §2.2.3, (26) is unacceptable for the same reason: there is no parallelism between the *than*-expression, *than Otto*, and the matrix clause, *a greater man would be a better man*. In contrast, in interpreting (29) (i.e., *he is a greater and better man than I am ~~a d-great and d′-good man~~*), parallelism is available.

2.2.5 Multi-head comparatives that have no overt *than*-clause

Sometimes multi-head comparatives are like regular comparatives in containing no overt *than*-clause (see (39) and examples like (5), (6), (8), (20), (24), and (25)).

(39) a. (Mary is exactly 6 feet tall.) Sue is taller. **Regular comparative**

b. Newer generations of microchips contain <u>more</u> electronic switches on a <u>smaller</u> surface. **Multi-head comparative** (Hendriks 1994: (14))

Evidently, in conducting a comparison with the comparison target (e.g., Sue's height in (39a)), the standard (e.g., Mary's height in (39a)) can be either (i) expressed in the same sentence by a *than*-clause/phrase or (ii) a discourse-salient item in the context.

Comparatives without an overt *than*-clause are called **discourse comparatives** in Hendriks (1994). Hendriks (1994) suggests that discourse comparatives are more interpretable than multi-head comparatives with an overt *than*-clause and further claims that 'comparatives may contain at most one instance of *sentence-internal* comparison (Hendriks 1994: (17))'. Here *sentence-internal* comparison means comparison based on an overt *than*-clause.

Examples like (2)/(13) clearly argue against this claim. In (2)/(13), the *than*-clause, *than cats ate mice*, supports the so-called sentence-internal comparison along both the dimensions of agent and theme cardinality, i.e., there are two comparisons.

However, it is likely that without an overt *than*-clause, parallelism between targets and standards involved in multi-dimensional comparison can often be easily accommodated, while sometimes, an overt *than*-clause blocks the establishment of parallelism (see (33) and (36)).

The role of comparative expressions (especially the meaning of comparative morpheme *-er*/*more*) as well as how to interpret comparatives with or without an overt *than*-clause will be discussed in greater detail in §4.1.

2.2.6 Pseudo-multi-head comparatives that conduct only one comparison

According to Hendriks (1994), since (i) multi-head comparatives like (16) suffer from the issue of mutual dependency and infinite regress (see §2.2.1) and (ii) 'comparatives may contain at most one instance of sentence-internal comparison (Hendriks 1994: (17))', a multi-head comparative with an overt *than*-clause actually conducts only one comparison, as illustrated in (40).

(40) John made <u>more</u> people prettier <u>than</u> I thought he would. (Hendriks 1994: (12))
⇝ John made <u>more</u> people prettier <u>than</u> I thought he would ~~make m-many people prettier~~. (Hendriks 1994: (16))

In (40), the *than*-clause provides a measurement on how many people I thought he would make prettier. The matrix clause also provides a measurement on how many people John (actually) made prettier. There is only a comparison of human cardinalities

Similarly, in our most natural interpretation of a sentence like (9a) (repeated here as (41)), both the matrix and the *than*-clause provide information that measures how many schools choose a longer school year now / before.

(41) <u>More</u> schools choose a longer school year <u>than</u> before. (= (9a))
⇝ More schools choose a longer school year than ~~m-many schools chose a longer school year~~ before

(40) and (9a)/(41) should not be considered genuine multi-head comparatives. Genuine multi-head comparatives like (2) or (3) involve multiple comparisons along different dimensions simultaneously, and their matrix and *than*-clause provide parallel multi-dimensional measurements, not single-dimensional measurements (see (38)).

Can the matrix and the *than*-clause of (40)/(41) provide parallel multi-dimensional measurements that correspond to multi-degree questions (42)/(43)? Similarly, can the matrix and the *than*-clause of (27) (see §2.2.3) resolve a multi-degree question like (44)?

(42) How many people did John make how pretty? ($\rightsquigarrow$ (40))

(43) How many schools choose how long a school year? ($\rightsquigarrow$ (41))

(44) *More doors are higher than windows are wide. (= (27))
$\rightsquigarrow$ How many doors/windows are how high/wide?

Intuitively, (42)–(44) are not syntactically unacceptable. However, as already noted by Meier (2001), the interpretability and acceptance of a multi-head comparative highly depends on whether we can naturally imagine a situation in which the multi-head comparative sounds relevant.

Whether multi-degree questions (42)–(44) make sense depends on whether we can imagine a situation in which raising these questions sounds meaningful.

(42) might be relevant in a context where interlocutors are interested in the overall performance of John, a plastic surgeon, and this overall performance is based on both the measurements of quantity (i.e., how many people) and quality (i.e., how pretty).

For (43) and (44), the two measurements involved lack a natural connection, i.e., it is difficult to imagine a context where the information of the two degree expressions can be integrated holistically to make sense. For (43), given that there is no underlying concern that depends on the number of schools and the length of a school year simultaneously, (43) is pragmatically weird, and thus (9a)/(41) is most naturally interpreted as conducting only one comparison.

2.3 Interim summary

To sum up, this section has shown empirically what characterizes a genuine multi-head comparative. A genuine multi-head comparative conducts multiple comparisons along different dimensions simultaneously (§2.1).

These multi-dimensional comparisons can have the same or opposite directions (§2.2.1), demonstrate flexibility with measurement units (§2.2.2), and are not restricted to comparisons of quantity (cf. the 'plural/mass nouns' restriction in §2.2.3). Comparison standards involved in multi-dimensional comparisons can be expressed by an overt *than*-clause or inferred from context (§2.2.5).

Crucially, a genuine multi-head comparative requires parallelism between comparison standards (from the *than*-clause or context) and comparison targets (from the matrix clause) in providing multi-dimensional measurements (§2.2.4). Moreover, these multi-dimensional measurements need to be able to be integrated holistically, contributing to the resolution of a single issue in a context (§2.2.6).

The simultaneous multi-dimensional measurements and their underlying interplay to address a single issue are reminiscent of cumulative-reading sentences. Below I show how the analysis of multi-head comparatives can be developed based on existing understanding of cumulative-reading sentences.

3 Inspiration from cumulative-reading sentences

I first present Brasoveanu (2013)'s analysis of cumulative-reading sentences, focusing on the simultaneity of multi-dimensional measurements (§3.1). Then I follow Krifka (1999) and Zhang (2023)'s discussion to show the pragmatics involved in the interpretation of cumulative-reading sentences (§3.2): the interplay between multi-dimensional measurements addresses a single underlying issue. Based on these, an informal account for multi-head comparatives is presented in §3.3.

3.1 Brasoveanu (2013)'s analysis of cumulative-reading sentences

A sentence like (45) contains two modified numerals (see the underlined parts) and has both a **distributive** and a **cumulative** reading (see (45a) and (45b)).

(45) Exactly three boys saw exactly five movies. (see e.g., Brasoveanu 2013)

a. **Distributive**: There are exactly 3 boys such that each of them saw exactly 5 movies.

b. **Cumulative**: The total number of boys who saw any movies is 3, and the total number of movies seen by any boys is 5.

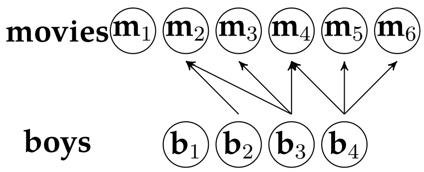

Figure 1: The genuine **cumulative** reading of (45) is **true** in this context.

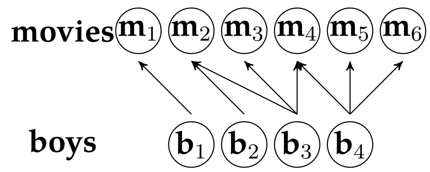

Figure 2: The genuine **cumulative** reading of (45) is **false** in this context.

Krifka (1999) and Brasoveanu (2013) emphasize that the derivation of the genuine cumulative reading (45b) is necessarily built on a **mutual restriction** between the two modified numerals, so that the two modified numerals are interpreted simultaneously and no scope-taking is involved.

As illustrated in (46), if one modified numeral (here *exactly three boys*) takes scope over the other (here *exactly five movies*), then only the latter is involved in restricting and defining the former, but not vice versa. The reading derived in (46) (dubbed the pseudo-cumulative reading) is true under the scenario in Fig. 2.

(46) **Unattested pseudo-cumulative reading** of (45): The maximal boy-sum such that 'they saw in total five movies between them' has a cardinality of 3.
$\rightsquigarrow$ True under the scenario in Fig. 2: there are two boy-sum witnesses, $\mathbf{b}_2 \oplus \mathbf{b}_3 \oplus \mathbf{b}_4$ and $\mathbf{b}_1 \oplus \mathbf{b}_2 \oplus \mathbf{b}_4$. There is no larger boy-sum satisfying the restriction that 'they saw in total five movies between them'.

Intuitively, sentence (45) is false under the scenario in Fig. 2, indicating that the pseudo-cumulative reading in (46) is too weak and unattested. Thus the genuine cumulative reading of (45) must be stronger and involve mutual restriction.

The genuine cumulative reading (see (45b)) is true under the scenario in Fig. 1. For this reading, *exactly three boys* denotes and counts the totality of **boys who saw any movies**, which is 3 (cf. the cardinality of all boys in the context is 4!), and *exactly five movies* denotes and counts the totality of **movies seen by any boys**, which is 5 (cf. the cardinality of all movies in the context is 6!).

Brasoveanu (2013) adopts dynamic semantics to implement the above ideas and develops a formal analysis for the cumulative reading of (45). Essentially, modified numerals make semantic contributions in several layers, as sketched in (47):

(47) Exactly three boys saw exactly five movies. **Cumulative reading** of (45)

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

First, each modified numeral introduces a potentially plural discourse referent (dref). Then restrictions are added onto drefs: here $\textsc{boy}(x)$, $\textsc{movie}(y)$, and $\textsc{saw}(x,y)$.[3]

Second, at the sentence level, after all the drefs are introduced and restrictions are added, modified numerals contribute **mereology-based maximality operators**. These maximality operators are applied at the sentence level simultaneously, picking out the maximal drefs satisfying the restrictions $\textsc{boy}(x)$, $\textsc{movie}(y)$, and $\textsc{saw}(x,y)$.

Finally, the modified numerals contribute **cardinality tests**, checking whether the cardinality of the relativized maximal boy-sum (i.e., those who saw movies) is 3 and whether the cardinality of the relativized maximal movie-sum is 5.

Evidently, there is a great similarity between cumulative-reading sentences and multi-head comparatives like (2) (see von Stechow 1984's analysis in (13)). Our intuitive interpretation for both of them involves mutual restriction and relativized maximization.

Mutual restriction paves the ground for the selection / definition of relativized maximal plural individuals. A cumulative-reading / multi-measurement sentence like (48a) **measures the cardinality of each relativized maximal plural individual**, while a multi-head comparative like (48b) **compares their cardinalities**.

(48) **Measurement and comparison along two dimensions: #agent and #theme**

a. Exactly three cats ate exactly five mice.
$\rightsquigarrow$ how many cats ate how many mice?

b. $\underbrace{\text{More dogs ate more rats } \underbrace{\text{than cats ate mice}}_{\text{how many cats ate how many mice}}}_{\text{how many dogs ate how many rats}}$ (= (2)/(13))

Thus the connection between multi-measurement sentence (48a) and multi-head comparative (48b) is parallel with that between measurement sentence (49a) and comparative (49b). For regular measurement sentences and comparatives, measurement and comparison are conducted along one single dimension (e.g., the cardinality of theme in (49)). For multi-measurement sentences and multi-head comparatives, measurement and comparison are conducted simultaneously along multiple dimensions (e.g., the cardinalities of agent and theme in (48)). Multi-head comparatives like (48b)

[3]Cumulative closure is assumed for lexical relations when needed.

are thus the comparative form of multi-measurement sentences.

(49) **Measurement and comparison along one dimension: the cardinality of theme**

a. Mary bought 6 books. $\rightsquigarrow$ How many books did Mary buy?

b. $\underbrace{\text{Sue bought more books } \underbrace{\text{than Mary did } \cancel{\text{buy } m\text{-many books}}}_{\text{how many books Mary bought}}}_{\text{how many books Sue bought}}$

3.2 Interplay between multi-dimensional measurements

Krifka (1999) discusses cumulative-reading sentences like (50), arguing that the **simultaneous mereology-based maximization** strategy (that works for the analysis of cumulative-reading sentences like (45)) cannot work for (50):

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

Krifka (1999) points out that

> 'this strategy would lead us to select the alternative *In Guatemala, 100 percent of the population own 100 percent of the land*, which clearly is not the most informative one among the alternatives – as a matter of fact, it is pretty uninformative.'
>
> 'What is peculiar with sentences like (50) 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.'
>
> (Krifka 1999: §3.1)

Krifka (1999)'s discussion shows that it is not numeric values alone that determine informativeness. For (50), increasing the number values (here '$\leq 3\%$' and '$\geq 70\%$') to higher alternatives (e.g., 100% and 100%) will not lead to higher informativeness, because here informativeness should be about a 'biased distribution'.

Fintel et al. (2014) makes related observations and discusses degree properties that are **upward-monotonic** vs. **downward-monotonic**. According to Fintel et al. (2014), for an upward-monotonic degree property like (51a), increasing the number values leads to higher informativeness. In contrast, for a downward-monotonic degree property like (51b), decreasing the number values leads to higher informativeness.

(51) a. **Degree properties that are upward-monotonic** (Fintel et al. 2014):
$\lambda n.$ Miranda has n kids
e.g., Miranda has 4 kids $>_{\textsc{info}}$ Miranda has 3 kids
('Miranda has 4 kids' entails 'Miranda has 3 kids')

b. **Degree properties that are downward-monotonic** (Fintel et al. 2014):
$\lambda d.d$-much walnuts is sufficient to make a pan of baklava
e.g., 300 grams of walnuts are sufficient to make a pan of baklava $>_{\textsc{info}}$ 500 grams of walnuts are sufficient to make a plan of baklava
('300 grams of walnuts are sufficient to make a pan of baklava' entails '500 grams of walnuts are sufficient to make a plan of baklava')

Zhang (2023) further points out that there should be a distinction between the notion of **logical entailment** and **informativeness**.[4]

As illustrated in (52), for a degree property like (51a), increasing or decreasing the number values does not always guarantee logical entailment.

(52) a. Miranda has exactly 4 kids $\nvDash$ Miranda has exactly 3 kids
b. Miranda has exactly 3 kids $\nvDash$ Miranda has exactly 4 kids

Nevertheless *Miranda has exactly 4 kids* corresponds to a higher degree in a context that addresses, e.g., the burden level of raising kids (see (53)). Here informativeness is not directly based on logical entailment that involves a **context-independent numerical scale**, but rather entailment that involves a **context-dependent scale of reached manyness**. '|kids| = 4' is guaranteed to reach a manyness degree that '|kids| = 3' does.

(53) Context: The more kids one has, the higher the burden level is.
Miranda has exactly 4 kids $>_{\textsc{reached-manyness}}$ Miranda has exactly 3 kids
$\rightsquigarrow$ '|kids| = 4' corresponds to a degree of manyness stronger than '|kids| = 3'

This kind of informativeness does not semantically depend on logical entailment (see (51a)), but rather pragmatically depends on a context-dependent degree QUD that addresses interlocutors' underlying concern related to their conversation goal.

In a context where people care about, e.g., the burden level, increasing kids' numbers (from 3 to 4) leads to a higher degree of burden, i.e., higher informativeness.

[4] Fintel et al. (2014) also mentioned in their footnote 4 that it is likely that some notion of informativity other than logical entailment is needed.

But in a context where one is qualified for a fixed amount of stipend if one has at least 3 kids, *Miranda has exactly 4 kids* is not considered more informative than *Miranda has exactly 3 kids* in addressing the issue of how qualified she is for receiving this stipend.

Similarly, for cumulative-reading or multi-measurement sentences (see (45) and (50)), informativeness depends on an underlying, context-dependent issue. For (50), as pointed out by Krifka (1999), it 'gives information about the bias of a statistical distribution'. It is this kind of underlying issue that determines how the increase or decrease of number values leads to higher or lower informativeness.

Generally, in interpreting cumulative-reading or multi-measurement sentences, context leads to two kinds of interplays among multi-dimensional measurements:

1. **Distance-based maximal informativeness** is reached by maximizing / minimizing the measurements along all dimensions (see (45) and Fig. 3). Changing the values in the same direction along all dimensions leads to higher informativeness.

2. **Angle-based maximal informativeness** is reached by maximizing the ratio between two measurements (see (50) and Fig. 4). Here maximal informativeness is reached when one measurement is high but the other is low.

Thus a multi-measurement sentence addresses a single context-dependent underlying issue, which determines how maximal informativeness is calculated from the two measurements: as distance-based (see (45)) or angle/ratio-based (see (50)).

Oftentimes, this underlying issue is not overtly expressed or specified, but inferred from context and our world knowledge. A sentence like (45) might address issues like *how high the overall film consumption among boys is*,[5] while a sentence like (50), as pointed out by Krifka (1999), addresses 'particularly biased distributions'.

Sometimes how maximal informativeness is calculated is reflected in the monotonicity of modified numerals used in a multi-measurement sentence. For distance-based informativeness maximality (e.g., (45)), the two numerals contribute to informativeness in a parallel way, and thus they usually share the same monotonicity. For the calculation of maximal angle / ratio (e.g., (50)), the two numerals contribute to the informativeness in opposite ways, and thus they often show opposite monotonicity. In (50), *at most 3%* is downward-entailing, while *at least 70%* is upward-entailing.

[5] In Brasoveanu (2013)'s analysis of (45) (see (47)), the simultaneous application of mereology-based maximality operators actually reflects the implicit assumption of this kind of degree QUD, which involves a distance-based maximal informativeness.

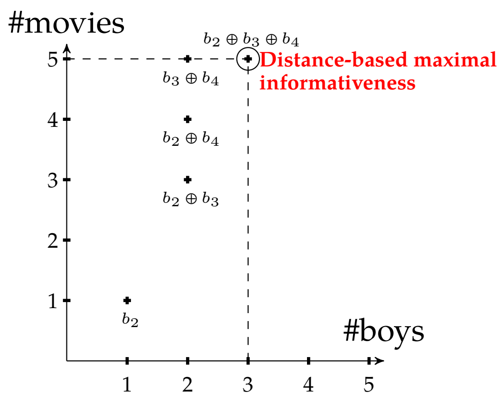

Figure 3: (45): *Exactly three boys saw exactly five movies*.
The plotting of the cardinalities of boy-sums and the cardinalities of movies they saw is shown.
Maximal informativeness is represented by the right-uppermost dot (i.e., the one farthest to the origin point of the coordinate system), which corresponds to the boy-sum $b_2 \oplus b_3 \oplus b_4$ and the 5 movies they saw between them (see Fig. 1).

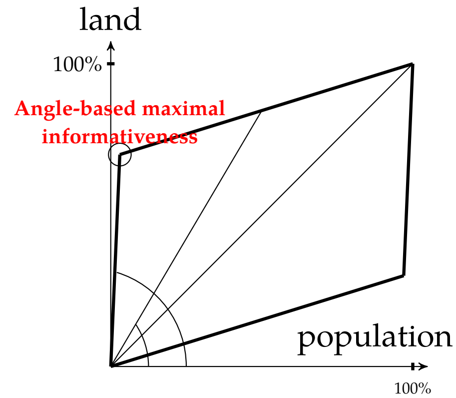

Figure 4: (50): *At most 3% of the population own* (*at least*) *70% of the land*.
The plotting of the percentages of land-owning populations and their total owned land is shown as a parallelogram-like area.
Maximal informativeness is represented by the left-uppermost corner (i.e., the one leading to the largest angle, or the largest ratio between two measurements), which corresponds to 'at most 3% of the population own at least 70% of the land'.

3.3 From multi-measurement sentences to multi-head comparatives

A multi-head comparative is the comparative form of a corresponding multi-measurement sentence. Thus, like its corresponding multi-measurement sentence, pragmatically, a multi-head comparative uses multi-dimensional comparisons to address a single context-dependent underlying issue.

For a regular comparative (see (54)), its matrix and *than*-clause provide parallel single-dimensional measurements. Pragmatically, the measurement value associated with the matrix clause addresses an underlying issue with higher informativeness than the measurement value associated with the *than*-clause.

(54) **From single-dimensional measurement to single-dimensional comparison**:
Betty has exactly 3 kids. Miranda has more kids (than Betty ~~has *n*-many kids~~).

How many kids Betty has — How many kids Betty has

How many kids Miranda has
⇝ Miranda's measurement has higher informativeness than Betty's does

For multi-head comparatives (see (55) and (56)), their matrix and *than*-clause provide parallel multi-dimensional measurements. Pragmatically, the multi-dimensional measurements associated with the matrix clause address an underlying issue with higher informativeness than the multi-dimensional measurements associated with the *than*-clause (see Fig. 5 and 6).

(55) **Multi-dimensional comparison (in a context with distance-based maximal informativeness)**:

$\underbrace{\text{More dogs ate more rats } \underbrace{\text{than } \cancel{m\text{-many}} \text{ cats ate } \cancel{n\text{-many}} \text{ mice.}}_{\text{How many cats ate how many mice}}}_{\text{How many dogs ate how many rats}}$ (=(2))

$\rightsquigarrow$ the cardinalities at the matrix level are associated with higher informativeness than those at the *than*-clause level

(56) **Multi-dimensional comparison (in a context with angle-based maximal informativeness)**:

$\underbrace{\text{Fewer people own more wealth here } \underbrace{\text{than } \cancel{m\% \text{ of the population own } n\% \text{ of the wealth}} \text{ elsewhere.}}_{\text{How many people own how much wealth elsewhere}}}_{\text{How many people own how much wealth here}}$ (=(12))

$\rightsquigarrow$ the measurements at the matrix level are associated with higher informativeness than those at the *than*-clause level

For multi-head comparative (2)/(55) (see Fig. 5), the calculation of informativeness in addressing the context-dependent underlying issue is distance-based. Thus the two comparative expressions (here *more dogs* and *more rats*) are of the same direction. Eventually, the whole multi-head comparative indicates an increase of informativeness (e.g., about how successful predators are) that is represented as an increase of distance.

Then, for a multi-head comparative like (12)/(56) (see Fig. 6), the calculation of informativeness in addressing the underlying issue is angle-based (i.e., ratio-based). Thus the two comparative expressions (here *fewer people* and *more of the overall wealth*) have opposite directions. Eventually, the whole multi-head comparative indicates an increase of informativeness (e.g., about the level of skewness in wealth distribution) that is represented as an increase of angle.

In this sense, compared with multi-measurement sentences (see the last two paragraphs in §3.2), multi-head comparatives give clearer clues on how maximal informativeness is calculated based on an underlying context-relevant issue. Comparative expressions with the same direction (see (2)) address a context issue with

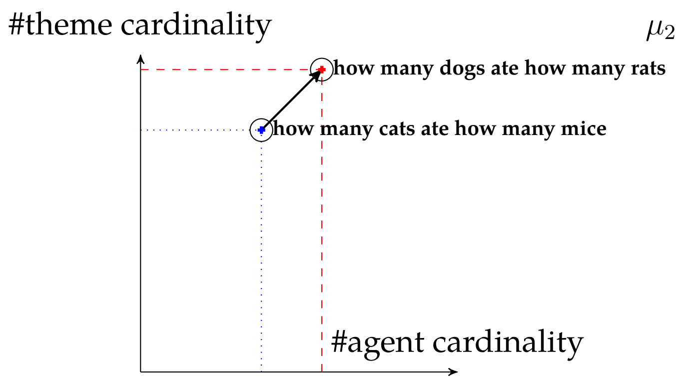

Figure 5: (2): *More dogs ate more rats than cats ate mice*.
The matrix and the *than*-clause address an underlying issue that leads to **distance-based** maximal informativeness.
The dot that represents the multi-dimensional measurements associated with the matrix clause has a farther distance from the origin than the dot that represents the multi-dimensional measurements associated with the *than*-clause. This increase of distance indicates an increase of informativeness in addressing the underlying issue in the context (e.g., the degree of success).

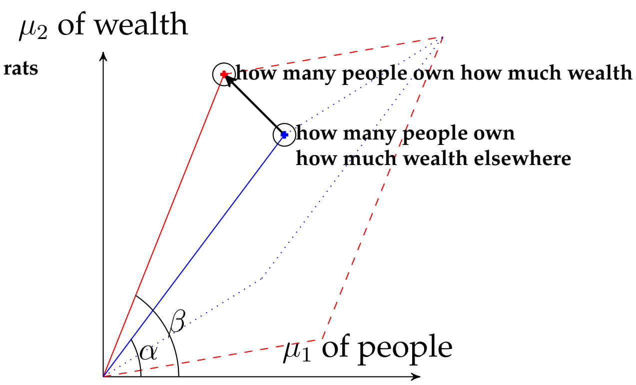

Figure 6: (12): *Fewer people own more of the overall wealth here than elsewhere*.
The matrix and the *than*-clause address an underlying issue that leads to **angle-based** maximal informativeness.
The dot that represents the multi-dimensional measurements associated with the matrix clause leads to a larger angle (here β) than the dot that represents the multi-dimensional measurements associated with the *than*-clause (see α). This increase of angle indicates an increase of informativeness in addressing the underlying issue in the context (e.g., the degree of skewness).

distance-based maximal informativeness, while comparative expressions with opposite directions (see (12)) address an issue with angle-based maximal informativeness.

Sometimes, an underlying context-relevant issue can also be explicitly expressed (e.g., *increased productivity* in (5)). In (57), a Portuguese example from Marques (2005), the underlying issue is about the quality of running, and this issue is resolved by the multi-dimensional measurements / comparisons along the scales of speed and risk frequency. The degree of running quality is calculated as the ratio between these two measurements.

(57) Correr bem tecnicamente é correr **mais** depressa com **menos** riscos.
to-run well technically is to-run **more** fast with **less** risks
'Technically, to run well is to run faster with less risks.' (Marques 2005: p.19, (52), from a corpus available at http://www.linguateca.pt/ACDC.)

To sum up, the ideas presented here capture all the crucial observations presented in §2. In a multi-head comparative, both the matrix and *than*-clause provide multi-dimensional measurements, supporting multi-dimensional comparisons between the matrix and the *than*-clause (see the discussion on multiple comparisons and parallelism in §2.1 and §2.2.4). Eventually, multi-dimensional comparisons contribute to resolve a single context-dependent issue (see §2.2.6), and whether multi-dimensional comparisons have the same or opposite directions (see §2.2.1) reflect how informativeness is calculated from measurement values in addressing this underlying context-dependent issue.[6] Based on these ideas, §4 will give a formal analysis.

4 Formal analysis of multi-head comparatives

The formal analysis of a multi-head comparative (§4.3) is developed by combining (i) an additivity-based view of comparatives implemented in interval semantics (§4.1) and (ii) a dynamic semantics implementation of maximizing informativeness (§4.2).

4.1 Additivity-based view of comparatives

4.1.1 Key insights of the additivity-base view and their motivation

According to the canonical view (see e.g., von Stechow 1984, Heim 1985, Kennedy 1999, Schwarzschild 2008, Beck 2011), (i) English comparative morpheme *-er*/*more* essentially performs comparison by encoding a '>' relation between two scalar values represented as degrees, and consequently, (ii) a comparative expresses an inequality (see (58)).

(58) **Canonical view: *-er* encodes an inequality relation: '>'**

Sue is taller than Mary is ~~tall~~. $\textsc{height}(\text{Sue}) > \textsc{height}(\text{Mary})$

LF: [-er [λd.Mary is d-tall]] [$\lambda d'$.Sue is d'-tall]

(Degree abstraction at both the matrix and the *than*-clause level)

a. $[\![\text{tall}]\!]_{\langle d,et\rangle} \stackrel{\text{def}}{=} \lambda d.\lambda x.\textsc{height}(x) \geq d$ a relation between d and x

$\rightsquigarrow$ the height of x reaches the degree d, i.e., x is tall to degree d

b. $[\![\text{-er}]\!]_{\langle\langle dt\rangle,\langle dt,t\rangle\rangle} \stackrel{\text{def}}{=} \lambda D_1.\lambda D_2.\textsc{max}(D_2) > \textsc{max}(D_1)$

$(\textsc{max} \stackrel{\text{def}}{=} \lambda D.\iota d[d \in D \wedge \forall d'[d' \in D \rightarrow d' \leq d]])$ (see e.g., Beck (2011))

[6] The rest of the discussion in §2 is about how multi-head comparatives are similar to regular comparatives in (i) being flexible with measurement units (see §2.2.2), (ii) supporting measurements / comparisons beyond quantity (see §2.2.3), and (iii) not necessarily requiring the overt presence of a *than*-clause (see §2.2.5).

Specifically, under the canonical view, a gradable adjective relates an individual x and a degree d, meaning that the measurement of x reaches d along a relevant scale (see (58a)). With degree abstraction, 'λd.Mary is d-tall' means the set of all height degrees the height of Mary reaches. The LF of a comparative assumes degree abstraction at both the matrix and the *than*-clause level, leading to two sets of degrees: those reached by Mary's height and those reached by Sue's height. Comparative morpheme *-er* relates and compares the maximal degree of the two sets of degrees (see (58b)).

Under the additivity-based view (see Zhang and Ling 2021, Zhang and Zhang 2024), (i) English comparative morpheme *-er/more* is an additive particle similar to *another*, denoting an increase that is anaphoric to a base item, and (ii) a comparative expresses an increase (or decrease) along a relevant scale.

(59) **Additivity-based view of comparatives**

Sue is taller than Mary is ~~tall~~.

$\rightsquigarrow$ Along a scale of height (or tallness reached), HEIGHT(Sue) (at the matrix level) includes an increase on the base item, HEIGHT(Mary) (at the *than*-clause level).

The parallelism between *-er/more* and additive particle *another* is inspired by the **additive use of *-er/more*** (see also Greenberg 2010, Thomas 2010, 2011 for discussion).

(60) **Parallelism between *another* and additive *more* in a domain of entities**

a. I ate $\underbrace{\text{an apple}}_{\text{the base}}$. Then I ate $\underbrace{\textbf{another}}_{\text{an increase}}$.

b. I ate $\underbrace{\text{two bars of chocolate}}_{\text{the base}}$. Then I ate $\underbrace{\text{(a bit) } \textbf{more} \text{ (chocalate)}}_{\text{an increase}}$.

In (60), both *another* (see (60a)) and *more* (see (60b)) denote an increase on a base item in a domain of **entities**. For (60b), the quantity of '(*a bit*) *more* (*chocolate*)' does not necessarily exceed the quantity of '*two bars*' in the precedent sentence, indicating that here a non-comparative interpretation of *more* exists, i.e., the additive use.

The **comparative use of *-er/more*** can also be considered an increase on a base item, but in domains of **scalar values** (i.e., items like **degrees** or **intervals** – convex sets of degrees). In (61a) and (61b), the base value provides information about the height of Mary, and *-er* denotes an increase on this base value. This increase can be further measured, as illustrated by the numerical restriction *2 inches* in (61a) and (61b).

(61) **Comparative use: the increase and its base are in domains of scalar values**

a. Mary is tall. Sue is $\underbrace{\text{(2 inches)}}_{\text{the restriction of the increase}}$ tall $\underbrace{\textbf{er}}_{\text{an increase based on the height of Mary}}$.

b. Sue is $\underbrace{\text{(2 inches)}}_{\text{the restriction of the increase}}$ tall $\underbrace{\textbf{er}}_{\text{an increase}}$ $\underbrace{\text{than Mary is } \cancel{\text{tall}}}_{\text{the base: the height of Mary}}$.

The comparative use of *-er* in (61b) is also parallel to additive particle *another* in (62) in (i) occurring in the same sentence (cf. the cross-sentential use of *another*/*more*/*-er* in (60) and (61a)) and (ii) being restricted.

(62) $\underbrace{\text{A girl, Mary}}_{\text{the base and its restriction}}$, met $\underbrace{\textbf{another}\text{ girl, Sue}}_{\text{an increase and its restriction}}$.

(63) provides further evidence on the parallelism between the comparative and additive uses of *-er*/*more* and *another*. They can all be used repetitively, expressing the accumulation of increases (see Zhang and Zhang 2024).

(63) **Repetitive use of *-er*/*more* and *another*: the accumulation of increases**

a. Lucy is tall<u>er</u> and tall<u>er</u>. **Comparative**

b. I bought <u>more</u> and <u>more</u> books. **Additive**

c. Janice had a little lamb and <u>another</u> and <u>another</u> and <u>another</u>. **Additive**

Since the use of additive particles like *another* and *-er*/*more* requires a discourse-salient item serving as the base, *another* and *-er*/*more* are anaphoric items.

Beaver and Clark (2009) (see also Thomas 2011's discussion on *another*) points out that **additivity** should be considered a phenomenon of **QUD-based anaphoricity** (see also Roberts 1996/2012, Büring 2003, Zeevat 2004, Zeevat and Jasinskaja 2007). Additive particles (e.g., *another*, *also*) are anaphoric to a contextually salient QUD, indicating an extension of a salient partial answer in addressing the current question.

To address a current degree QUD, the comparative use of *-er*/*more* requires a salient partial answer from context (see (61a)) or the *than*-clause (see (61b)/(64)/(65)) and extends this partial answer, marking a higher degree of informativeness.

(64) $\underbrace{\text{Sue is (2 inches) tall } \underbrace{\text{er}}_{\textbf{an increase}} \underbrace{\text{than Mary is } \cancel{\text{tall}}}_{\substack{\textbf{the base: }\text{Mary's height} \\ \textbf{A salient partial answer in addressing } \textbf{\textit{how tall Sue is}}\textbf{:} \\ \text{the degree of tallness that Mary's height reaches}}}}_{\substack{\text{Sue's height} \\ \textbf{Addressing the current degree QUD: } \textbf{\textit{how tall Sue is}} \\ \text{the degree of tallness that Sue's height reaches}}}$ (= (61b), see Fig. 7)

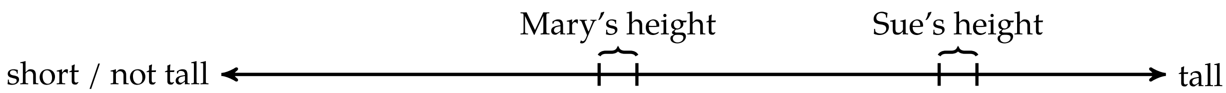

Figure 7: Semantically, a comparative expresses an increase / decrease to a scalar value. Pragmatically, it expresses an increase of degree-QUD-dependent informativeness. For a QUD on the degree of tallness reached (e.g., (64)), the increase of height leads to higher informativeness. For a QUD on the degree of shortness reached (e.g., (65)) or the degree of tallness not reached, the decrease of height leads to higher informativeness.

(65) Mary is (2 inches) short er than Sue is ~~short~~ (see Fig. 7)

er: **an increase**

than Sue is ~~short~~: **the base:** Sue's height
A salient partial answer in addressing *how short Mary is*:
the degree of shortness that Sue's height reaches

Mary is (2 inches) short er than Sue is ~~short~~: Mary's height
Addressing the current degree QUD: *how short Mary is*
the degree of shortness that Mary's height reaches

As illustrated in (64) and (65) (see also Fig. 7), semantically, comparatives encode an increase or decrease between context-independent measurements. In Fig. 7, these measurements are represented as (not-very-precise) interval positions along a scale.

Intuitively, (64) and (65) have the same truth condition, indicating that gradable adjectives like *tall* and *short* (see (64) and (65)) are associated with the same scale, but with different directions (see Fig. 7). For example, Mary's height is at the position of the interval $[5'11'', 6']$, independent from scale direction.

Pragmatically, the scalar value associated with the *than*-clause is a partial answer in addressing the current degree QUD at the matrix level. Thus, for a *taller-than* sentence (see (64)), the QUD is about the degree of tallness reached, and the matrix level is associated with higher informativeness, corresponding to a higher height measurement. Conversely, for a *shorter-than* (see (65)) or *less-tall-than* sentence, the QUD is about the degree of shortness reached or the degree of tallness not reached, and the matrix level is associated with higher informativeness, corresponding to a lower height measurement.

Compared with the canonical view, the additivity-based view on comparatives has at least two conceptual advantages. First, it provides a unified account for the additive and comparative use of English morpheme *-er*/*more*: *-er* denotes an increase (in a domain of entities or scalar values). Second, the additivity-based view analyzes *-er*/*more* and comparatives along with other additivity-related phenomena. They are all QUD-based and extend a salient partial answer to higher informativeness.

4.1.2 Formal analysis of comparatives

I implement the additivity-based view on comparatives in **interval semantics**, using **interval subtraction**, instead of inequalities, to formally implement comparison (see also Zhang and Ling (2021) and Zhang and Zhang (2024)).

0. Interval semantics We can use an interval to represent a not-very-precise position along a scale (see Fig. 7). An interval is a convex set of degrees. Given that the type of degrees is d, the type of intervals is $\langle dt \rangle$.

A totally ordered set S is convex if for any two items $a, b \in S$ (suppose $a < b$), then $\forall x[a \le x \le b \rightarrow x \in S]$. Thus intervals can be written with their upper and lower bounds: square brackets '[' and ']' mean closed lower and upper bounds, and round parentheses '(' and ')' mean open lower and upper bounds. E.g.,

(66) $\{x \mid I_{\text{min}} < x \le I_{\text{max}}\} = (I_{\text{min}}, I_{\text{max}}]$ A left-open and right-closed interval
$\{x \mid I_{\text{min}} \le x < I_{\text{max}}\} = [I_{\text{min}}, I_{\text{max}})$ A left-closed and right-open interval

A singleton set like $\{x \mid x = 6'1''\}$ can be written as $[6'1'', 6'1'']$. '$+\infty$' and '$-\infty$' mean positive and negative infinity. Thus $(0, +\infty)$ is the shorthand for $\{d \mid 0 < d < +\infty\}$.

The interval subtraction between two intervals I_1 and I_2 results in the largest range of possible differences between any two points in I_1 and I_2:

(67) $$\underbrace{[y_1, y_2]}_{\textbf{minuend: target's measurement}} - \underbrace{[x_1, x_2]}_{\textbf{subtrahend: standard}} = \underbrace{[y_1 - x_2, y_2 - x_1]}_{\textbf{difference: differential}}$$

Thus, given the subtrahend $[a, b]$ and the difference $[c, d]$,

(68) Minuend $= [b + c, a + d]$ (defined when $b + c \le a + d$)

See Moore (1979) for more details on interval notations and operations.

The use of interval semantics, instead of degree semantics, has two advantages. First, intervals are more generalized than degrees in representing scalar values, supporting the expression of not-very-precise values in natural language (e.g., *Mary is between 5 feet 11 inches and 6 feet tall*, see Fig. 7). Second, compared with number arithmetic, interval arithmetic (see (67) and (68)) provides a richer technique in showing formal properties of comparison (e.g., the monotonicity of the subtrahend (i.e., comparison standard, which is typically contributed by a *than*-clause/phrase), see Zhang 2020, Zhang and Ling 2021).

1. The semantics of gradable adjectives I follow Zhang and Zhang (2024) and analyze the semantics of a gradable adjective as a substraction relation among three scalar values, all represented as intervals:

(69) $$[\![\text{tall}]\!] \stackrel{\text{def}}{=} \lambda I_{\text{DIFF}}.\lambda I_{\text{STDD}}.\lambda x.\ \underbrace{I_{\text{DIFF}} \subseteq [0,+\infty)}_{\text{non-negative presupposition of the difference}}\ \textsc{height}(x) \subseteq \iota I[I - I_{\text{STDD}} = I_{\text{DIFF}}]$$

(i.e., along a height scale, $\textsc{height}(x)$ is compared to a reference position I_{STDD}, which is not higher than $\textsc{height}(x)$, and $\textsc{height}(x)$ is away from I_{STDD} by I_{DIFF})

(70) $$[\![\text{short}]\!] \stackrel{\text{def}}{=} \lambda I_{\text{DIFF}}.\lambda I_{\text{STDD}}.\lambda x.\ \underbrace{I_{\text{DIFF}} \subseteq [0,+\infty)}_{\text{non-negative presupposition of the difference}}\ \textsc{hight}(x) \subseteq \iota I[I_{\text{STDD}} - I = I_{\text{DIFF}}]$$

(i.e., along a height scale, $\textsc{height}(x)$ is compared to a reference position I_{STDD}, which is not lower than $\textsc{height}(x)$, and $\textsc{height}(x)$ is away from I_{STDD} by I_{DIFF})

Various uses of gradable adjectives can be immediately accounted for, as summarized in Table (71).

(71) **Various uses of gradable adjectives in English**

	Reference position I_{STDD}	Distance I_{DIFF}
Positive use	Contextual threshold: $[d^c_{\text{POS}}, d^c_{\text{POS}}]$	$[0,+\infty)$ (or further restricted by a modifier like *very*)
Measure constructions	Absolute zero point: $[0,0]$	restricted by a measure phrase
Degree questions	Contextual threshold or absolute zero point	(interval abstraction)

For the positive use of *tall*/*short*, $\textsc{height}(x)$ is compared to a relevant contextual threshold value for being tall/short (see (72) and (73)).

(72) $$\begin{aligned}[\![\text{Brienne is POS tall}]\!] &\Leftrightarrow \textsc{height}(\text{Brienne}) \subseteq \iota I[I - [d^c_{\text{POS}}, d^c_{\text{POS}}] = [0,+\infty)] \\ &\Leftrightarrow \textsc{height}(\text{Brienne}) \subseteq [d^c_{\text{POS}}, +\infty)\end{aligned}$$

(i.e., the height of Brienne reaches the contextual threshold of being tall)

(73) $$\begin{aligned}[\![\text{Tyrion is POS short}]\!] &\Leftrightarrow \textsc{height}(\text{Tyrion}) \subseteq \iota I[[d^c_{\text{POS}'}, d^c_{\text{POS}'}] - I = [0,+\infty)] \\ &\Leftrightarrow \textsc{height}(\text{Tyrion}) \subseteq (-\infty, d^c_{\text{POS}'}]\end{aligned}$$

(i.e., Tyrion's height is the same as or below the threshold of being short)

For the measurement use of *tall*, $\textsc{height}(x)$ is compared to the absolute zero point along a height scale, and the distance I_{DIFF} is restricted by a measure phrase (e.g., $[5'11'', 6']$ in (74)). Obviously, $\textsc{height}(x)$ cannot be lower than $[0, 0]$. Thus a sentence like (75) that involves the use of a gradable adjetive like *short* is unacceptable.

(74) $[\![$Mary is <u>between 5 feet 11 inches and 6 feet</u> tall$]\!]$
$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I - [0, 0] = [5'11'', 6'] \cap [0, +\infty)]$
$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq [5'11'', 6']$

(75) *Chloe is 5 feet short. $[0, 0] - \textsc{height}(\text{Chloe})$ is negative (see (70))
Context: Chloe's measurement falls at the position $[5', 5']$ on a height scale.

The derivation of a degree question involves abstraction with regard to I_{DIFF}, giving rise to a set of intervals that provides information on the distance between $\textsc{height}(x)$ and a reference position (see Hausser and Zaefferer 1978, Hausser 1983's categorial approach to questions). A reference position is typically the absolute zero point (see (76a)) or a contextual threshold (see (76b) and (77)). When the reference position is a contextual threshold, the degree question is interpreted with evaluativity (see (76b) and (77)).

(76) $[\![$How tall is Brienne$]\!] \Leftrightarrow \lambda I_{\text{DIFF}}.\textsc{height}(\text{Brienne}) \subseteq \iota I[I - I_{\text{STDD}} = I_{\text{DIFF}}]$
a. I_{STDD} is equal to $[0, 0]$: **No evaluativity**
$[\![(76)]\!] \Leftrightarrow \lambda I_{\text{DIFF}}.\textsc{height}(\text{Brienne}) \subseteq \iota I[I - [0, 0] = I_{\text{DIFF}}]$
(i.e., how far is Brienne's height away from the zero point)
b. I_{STDD} is equal to $[d^c_{\text{POS}}, d^c_{\text{POS}}]$: **Evaluativity**: $\rightsquigarrow$ Brienne is tall
$[\![(76)]\!] \Leftrightarrow \lambda I_{\text{DIFF}}.\textsc{height}(\text{Brienne}) \subseteq \iota I[I - [d^c_{\text{POS}}, d^c_{\text{POS}}] = I_{\text{DIFF}}]$
(i.e., to what extent is Brienne's height above the threshold of being tall)

(77) $[\![$How short is Tyrion$]\!] \Leftrightarrow \lambda I_{\text{DIFF}}.\textsc{height}(\text{Tyrion}) \subseteq \iota I[I_{\text{STDD}} - I = I_{\text{DIFF}}]$
I_{STDD} is equal to $[d^c_{\text{POS}'}, d^c_{\text{POS}'}]$: **Evaluativity**: $\rightsquigarrow$ Tyrion is short
$[\![(77)]\!] \Leftrightarrow \lambda I_{\text{DIFF}}.\textsc{height}(\text{Tyrion}) \subseteq \iota I[[d^c_{\text{POS}'}, d^c_{\text{POS}'}] - I = I_{\text{DIFF}}]$
(i.e., to what extent is Tyrion's height below the threshold of being short)

An $\mathbf{Ans}_{\text{DIFF}}$ picks out the most informative interval from a set (see Dayal 1996):

(78) An answerhood operator $\mathbf{Ans}_{\text{DIFF}}$ is defined for a set of intervals p s.t.
$\exists I[p(I) \wedge \forall I'[[p(I') \wedge I' \neq I] \rightarrow I \subset I']]$
When defined, $\mathbf{Ans}_{\text{DIFF}} \stackrel{\text{def}}{=} \lambda p_{\langle dt,t\rangle}.\iota I[p(I) \wedge \forall I'[[p(I') \wedge I' \neq I] \rightarrow I \subset I']]$

When $\mathbf{Ans}_{\text{DIFF}}$ is applied to a degree question, it returns the most informative interval that addresses how far the position representing the measurement of a target is away from a certain reference position.

2. The semantics of comparative morpheme *-er/more* Within an interval-theoretic framework, ⟦-er/more⟧ denotes the default, most general positive interval, $(0, +\infty)$ (see (79a)). A numerical differential can further restricts the size of this increase (see (79b)).

(79) a. $[\![\text{-er/more}]\!]_{\langle dt \rangle} \overset{\text{def}}{=} (0, +\infty)$ **As an increase, it requires a salient base value**
b. $[\![\text{2 inches} \ldots \text{-er}]\!]_{\langle dt \rangle} = (0, +\infty) \cap [2'', +\infty) = [2'', +\infty)$ **('at least' interpretation)**

Operator LITTLE takes a positive interval as input and returns its inverse:

(80) $[\![\textsc{little}]\!]_{\langle dt, dt \rangle} \overset{\text{def}}{=} \lambda I.[0, 0] - I$

Based on the application of LITTLE, we can turn an increase into a decrease (i.e., a negative increase). Thus ⟦less⟧ denotes the default, most general negative interval, $(-\infty, 0)$ (see (81a)). When a numerical differential is used in a *less-than* comparative, the numerical differential first restricts the default difference, ⟦-er/more⟧, and then the application of LITTLE turns the restricted increase into its inverse (see (81b)).

(81) a. $[\![\text{less}]\!]_{\langle dt \rangle} = [\![\textsc{little}]\!][\![\text{-er/more}]\!] = (-\infty, 0)$ **a decrease**
b. $[\![\text{2 inches} \ldots \text{less}]\!]_{\langle dt \rangle} = [\![\textsc{little}]\!][\![\text{2 inches} \ldots \text{-er}]\!] = [0, 0] - [2'', +\infty) = (-\infty, 2'']$

3. The semantics of comparatives Under the additivity-based view, a comparative expresses an increase to a salient base value, making the matrix-level measurement more informative than the *than*-clause-level value in addressing a degree QUD. By serving as this base, the *than*-clause is considered a partial answer to the degree QUD.

As proposed above, a degree question asks about the distance away from a reference position (see (76) and (77)). Thus, as illustrated in (82) (see Fig. 8), with regard to a certain reference position, I_{STDD}, the *than*-clause ⟦than Mary is ~~tall~~⟧ amounts to the distance between Mary's height and I_{STDD} (see (82a)), and this distance serves as a partial answer in addressing the tallness degree reached by Sue's height at the matrix level. In (82b), at the matrix level, the distance between Sue's height and this reference I_{STDD} is an increase to the distance provided by the *than*-clause.

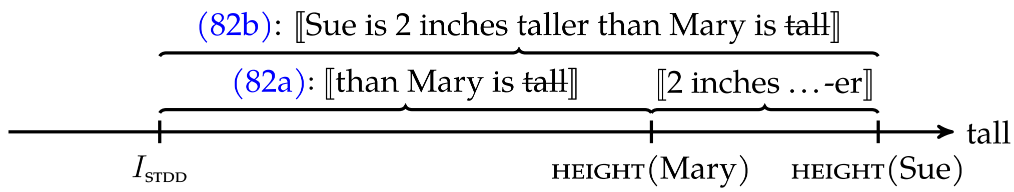

Figure 8: (82): *Sue is* (*2 inches*) *taller than Mary is ~~tall~~*. With an increase of distance away from a certain reference position, I_{STDD}, the matrix is more informative than the *than*-clause in addressing **the degree of tallness reached**.

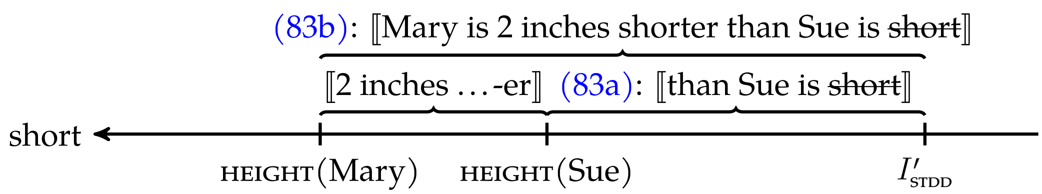

Figure 9: (83): *Mary is* (*2 inches*) *shorter than Sue is ~~short~~*. With an increase of distance away from a certain reference position, I'_{STDD}, the matrix is more informative than the *than*-clause in addressing **the degree of shortness reached**.

(82) $\underbrace{\text{Sue is (2 inches) tall } \underbrace{\text{er}}_{\textbf{an increase}} \ \underbrace{\text{than Mary is } \cancel{\text{tall}}}_{(82a)}}_{(82b)}$ (= (61b)/(64), see Fig. 8)

a. ⟦than Mary is ~~tall~~⟧

$\Leftrightarrow \textbf{Ans}_{\text{DIFF}}[\lambda I_{\text{DIFF}}.\textsc{height}(\text{Mary}) \subseteq \iota I[I - I_{\text{STDD}} = I_{\text{DIFF}}]]$

(i.e., the distance that Mary's height is above a reference position I_{STDD})

b. ⟦Sue is 2 inches taller than Mary is ~~tall~~⟧

$\Leftrightarrow \textsc{height}(\text{Sue}) \subseteq \iota I[I - I_{\text{STDD}} =$

$\underbrace{\iota I_{\text{DIFF}}[I_{\text{DIFF}} - \underbrace{[\![\text{than Mary is } \cancel{\text{tall}}]\!]}_{\text{(82a): the distance between Mary's height and } I_{\text{STDD}}} = (0, +\infty) \cap [2'', +\infty)]]}_{\text{an increase (of the size } [2'', +\infty)\text{) to the distance provided by the } \textit{than}\text{-clause}}$

The analysis of a *shorter-than* comparative is parallel to a *taller-than* comparative. As illustrated in (83) (see Fig. 9), the *than*-clause ⟦than Sue is ~~short~~⟧ provides information on the distance between Sue's height and a certain reference position, I'_{STDD} (see (83a)), and this distance serves as a partial answer in addressing the shortness degree reached by Mary's height. In (83b), at the matrix level, the distance between Mary's height and this reference I'_{STDD} is an increase to the distance provided by the *than*-clause.

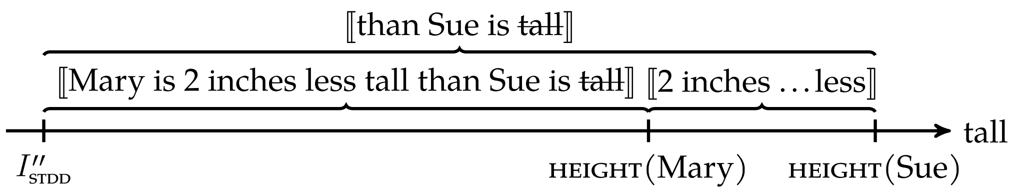

Figure 10: (84): *Mary is* (*2 inches*) *less tall than Sue is ~~tall~~*. The matrix is more informative than the *than*-clause in addressing **the degree of tallness NOT reached**.

(83) $\underbrace{\text{Mary is (2 inches) short } \underbrace{\text{er}}_{\textbf{an increase}} \underbrace{\text{than Sue is } \cancel{\text{short}}}_{(83a)}}_{(83b)}$ (= (65), see Fig. 9)

a. $[\![\text{than Sue is } \cancel{\text{short}}]\!]$

$\Leftrightarrow \textbf{Ans}_{\text{DIFF}}[\lambda I_{\text{DIFF}}.\textsc{height}(\text{Sue}) \subseteq \iota I[I'_{\text{STDD}} - I = I_{\text{DIFF}}]]$

(i.e., the distance that Sue's height is below a reference position I'_{STDD})

b. $[\![\text{Mary is 2 inches shorter than Sue is } \cancel{\text{short}}]\!]$

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I'_{\text{STDD}} - I =$

$\underbrace{\iota I_{\text{DIFF}}[I_{\text{DIFF}} - \underbrace{[\![\text{than Sue is } \cancel{\text{short}}]\!]}_{\text{(83a): the distance between Sue's height and } I'_{\text{STDD}}} = (0, +\infty) \cap [2'', +\infty)]]}_{\text{an increase (of the size } [2'', +\infty)) \text{ to the distance provided by the } \textit{than}\text{-clause}}$

The derivation of a *less-than* comparative is also similar. As illustrated in (84) (see Fig. 10), the *than*-clause ⟦than Sue is ~~tall~~⟧ provides information on the distance between Sue's height and a reference position, I''_{STDD} (see (84a)), and this distance serves as a partial answer in addressing the tallness degree not reached by Mary's height. In (84b), at the matrix level, the distance between Mary's height and this reference I''_{STDD} includes a decrease to the distance provided by the *than*-clause.

(84) $\underbrace{\text{Mary is } \underbrace{\text{(2 inches) less}}_{\textbf{a decrease}} \text{ tall } \underbrace{\text{than Sue is } \cancel{\text{tall}}}_{(84a)}}_{(84b)}$ (see Fig. 10)

a. $[\![\text{than Sue is } \cancel{\text{tall}}]\!]$

$\Leftrightarrow \textbf{Ans}_{\text{DIFF}}[\lambda I_{\text{DIFF}}.\textsc{height}(\text{Sue}) \subseteq \iota I[I - I''_{\text{STDD}} = I_{\text{DIFF}}]]$

(i.e., the distance that Sue's height is above a reference position I''_{STDD})

b. $[\![\text{Mary is 2 inches less tall than Sue is } \cancel{\text{tall}}]\!]$

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I - I''_{\text{STDD}} =$

$$\underbrace{\iota I_{\text{DIFF}}[I_{\text{DIFF}} - \underbrace{[\![\text{than Sue is } \cancel{\text{tall}}]\!]}_{\text{(84a): the distance between Sue's height and } I_{\text{STDD}}} = \underbrace{(-\infty, 2'']}_{\text{(81b)}}]]}_{\text{a decrease (of the size } [2'', +\infty)\text{) to the distance provided by the } \textit{than}\text{-clause}}$$

It is worth noting that in the interpretation of a comparative, a reference position like $I_{\text{STDD}}/I'_{\text{STDD}}/I''_{\text{STDD}}$ is not necessarily a contextual threshold for being tall or short, thus evaluativity (i.e., a positive inference, e.g., Mary is short) is never involved.

A reference position is a free variable. As shown in Fig. 8–10, its value does not matter in interpreting comparatives, as far as the reference keeps the same in deriving the meaning at the *than*-clause and matrix level. This is not an additional stipulation, because under the additivity-based view, both the matrix and the *than* clause address the same QUD and the *than*-clause works like a partial answer for a (positive or negative) increase to add on. Thus we can also set the reference position $I_{\text{STDD}}/I'_{\text{STDD}}/I''_{\text{STDD}}$ at special positions to simplify (82b), (83b), and (84b) as follows (see also Fig. 8–10):[7]

(85) $[\![\text{Sue is 2 inches taller than Mary is } \cancel{\text{tall}}]\!]$ $\qquad I_{\text{STDD}} = \textsc{height}(\text{Mary})$ in Fig. 8

$\Leftrightarrow \textsc{height}(\text{Sue}) \subseteq \iota I[I - I_{\text{STDD}} = \iota I_{\text{DIFF}}[I_{\text{DIFF}} - [\![\text{than Mary is } \cancel{\text{tall}}]\!] = [2'', +\infty)]]$ $\qquad$ (82b)

$\Leftrightarrow \textsc{height}(\text{Sue}) \subseteq \iota I[I - \textsc{height}(\text{Mary}) = \iota I_{\text{DIFF}}[I_{\text{DIFF}} - [0, 0] = [2'', +\infty)]]$

$\Leftrightarrow \textsc{height}(\text{Sue}) \subseteq \iota I[I - \textsc{height}(\text{Mary}) = [2'', +\infty)]$

(86) $[\![\text{Mary is 2 inches shorter than Sue is } \cancel{\text{short}}]\!]$ $\qquad I'_{\text{STDD}} = \textsc{height}(\text{Sue})$ in Fig. 9

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I'_{\text{STDD}} - I = \iota I_{\text{DIFF}}[I_{\text{DIFF}} - [\![\text{than Sue is } \cancel{\text{short}}]\!] = [2'', +\infty)]]$ (83b)

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[\textsc{height}(\text{Sue}) - I = \iota I_{\text{DIFF}}[I_{\text{DIFF}} - [0, 0] = [2'', +\infty)]]$

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[\textsc{height}(\text{Sue}) - I = [2'', +\infty)]$

(87) $[\![\text{Mary is 2 inches less tall than Sue is } \cancel{\text{tall}}]\!]$ $\qquad I''_{\text{STDD}} = \textsc{height}(\text{Sue})$ in Fig. 10

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I - I''_{\text{STDD}} = \iota I_{\text{DIFF}}[I_{\text{DIFF}} - [\![\text{than Sue is } \cancel{\text{tall}}]\!] = (-\infty, 2'']]]$ $\qquad$ (84b)

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I - \textsc{height}(\text{Sue}) = \iota I_{\text{DIFF}}[I_{\text{DIFF}} - [0, 0] = (-\infty, 2'']]]$

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[I - \textsc{height}(\text{Sue}) = (-\infty, 2'']]$

$\Leftrightarrow \textsc{height}(\text{Mary}) \subseteq \iota I[\textsc{height}(\text{Sue}) - I = [2'', +\infty)]$

Eventually, (86) and (87) amount to the same truth condition, but the analysis in (83) and (84) shows that they assume different comparisons and address different degree QUDs: i.e., the shortness degree reached vs. the tallness degree not reached.

[7] According to the lexical entries in (69) and (70), I_{DIFF} has a non-negative presupposition, so the simplification in (87) seems to violate this requirement. But mathematically, the simplification in (87) is totally fine, and the last two lines in (87) are equivalent.

4.2 The meaning of a cumulative-reading sentence in DPL

Following Brasoveanu (2013), I derive the meaning of cumulative-reading sentences within Dynamic Predicate Logic (DPL, Groenendijk and Stokhof 1991).

Within dynamic semantics, meaning derivation is considered a series of updates from an information state to another. Given the distributivity of DPL (see (88b), Groenendijk and Stokhof 1990), an update is a relation between assignment functions (i.e., it takes an assignment function as input and returns a set of assignment functions).

(88) DPL (see Groenendijk and Stokhof 1990, 1991):

a. **Information state** i: a set of assignment functions g — Type: $\langle st \rangle$

b. **Update** r: from an information state to another — Type: $\langle st, st \rangle$
Given that for every information state i, $r(i) = \bigcup_{g \in i} r(\{g\})$, an update can be considered a relation between assignment functions (of type $\langle s, st \rangle$).

c. **Truth**: an update is true if it does not end with an empty set.

According to Brasoveanu (2013) (see §3.1), modified numerals make semantic contribution in several layers: (i) introducing drefs (see (89a), where these drefs get restrictions like $\textsc{movie}(y), \textsc{boy}(x), \textsc{saw}(x, y)$); (ii) imposing mereology-based maximality operators $\mathbf{M}_{u,\nu}$ (see (89b), where the maximal boy-sum and movie-sum that satisfy all the relevant restrictions are picked out); (iii) imposing cardinality tests $\mathbf{3}_u$ and $\mathbf{5}_\nu$ (see (89c)). The cumulative reading of (45) is an update such that it is true if the cardinality of all boys who saw movies is 3 and the cardinality of all movies seen by boys is 5.

(89) $[\![$Exactly three^u boys saw exactly five$^\nu$ movies$]\!]$ (=(45)/(47))

$$\Leftrightarrow \underbrace{\mathbf{3}_u[\mathbf{5}_\nu}_{\text{cardinality tests}}\ [\ \underbrace{\mathbf{M}_{u,\nu}}_{\text{maximality}}\ \underbrace{[\![\text{some}^u \text{ boys saw some}^\nu \text{ movies}]\!]}_{\text{dref introduction}}]]$$

a. **Introducing drefs**: $[\![$some^u boys saw some$^\nu$ movies$]\!]$

$$\Leftrightarrow \lambda g\,.\left\{ g^{\nu \mapsto y}_{u \mapsto x} \,\middle|\, \textsc{movie}(y), \textsc{boy}(x), \textsc{saw}(x, y) \right\}$$

b. **Simultaneously applying mereology-based maximality operators**:

$$\mathbf{M}_{u,\nu} \overset{\text{def}}{=} \lambda m_{\langle s,st \rangle}.\lambda g_s.\{h \in m(g) \mid \neg \exists h' \in m(g).h(u) \sqsubset h'(u) \vee h(\nu) \sqsubset h'(\nu)\}$$

Thus $\mathbf{M}_{u,\nu}[\![$some^u boys saw some$^\nu$ movies$]\!]$

$$\Leftrightarrow \mathbf{M}_{u,\nu}[\lambda g\,.\left\{ g^{\nu \mapsto y}_{u \mapsto x} \,\middle|\, \textsc{movie}(y), \textsc{boy}(x), \textsc{saw}(x, y) \right\}]$$

$$\Leftrightarrow \lambda g\,.\left\{ g^{\nu \mapsto y}_{u \mapsto x} \,\middle|\, \begin{array}{l} y = \Sigma y[\textsc{movie}(y) \wedge \textsc{saw}(x, y)] \\ x = \Sigma x[\textsc{boy}(x) \wedge \textsc{saw}(x, y)] \end{array} \right\}$$

c. **Checking cardinalities**:

$$\mathbf{3}_u \stackrel{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } |g(u)| = 3 \\ \varnothing & \text{otherwise} \end{cases}$$

$$\mathbf{5}_\nu \stackrel{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } |g(\nu)| = 5 \\ \varnothing & \text{otherwise} \end{cases}$$

Thus $\mathbf{3}_u[\mathbf{5}_\nu[\mathbf{M}_{u,\nu}[\![\text{some}^u \text{ boys saw some}^\nu \text{ movies}]\!]]]$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x}} \,\middle|\, \begin{array}{l} y = \Sigma y[\textsc{movie}(y) \wedge \textsc{saw}(x,y)] \\ x = \Sigma x[\textsc{boy}(x) \wedge \textsc{saw}(x,y)] \end{array} \right\}, \text{ if } |x| = 3 \wedge |y| = 5$$

The discussion in §3.2 shows that Brasoveanu (2013)'s simultaneous application of two mereology-based maximality operators is a special case of a degree-QUD-based maximality of informativeness (see Fig. 3): maximal informativeness is achieved when measurements along both dimensions are maximal. Thus I define a generalized, degree-QUD-based maximality operator, which, among all drefs (assigned to $u_1, u_2, \ldots$), picks out those that lead to maximal informativeness in resolving a degree QUD:

(90) $\mathbf{M}_{\text{QUD-INFO}} \stackrel{\text{def}}{=} \lambda m.\lambda g.\{h \in m(g) \mid \neg\exists h' \in m(g).$
$G_{\text{QUD}}(\langle \mu_{11}(h'(u_1)), \mu_{12}(h'(u_1)), \ldots \mu_{21}(h'(u_2)), \ldots\rangle) >_{\text{INFO}}$
$G_{\text{QUD}}(\langle \mu_{11}(h(u_1)), \mu_{12}(h(u_1)), \ldots \mu_{21}(h(u_2)), \ldots\rangle)\}$

In (90), h, h' are assignment functions; $h(u_i), h(u_j), \ldots$ are drefs assigned to $u_i, u_j, \ldots$; μ_{ij} means a measure function for the dref assigned to u_i, mapping the dref to a value along a relevant scale. The same dref can be potentially measured by more than one measure functions, along multiple dimensions.

The definition of $\mathbf{M}_{\text{QUD-INFO}}$ in (90) includes a measure function G_{QUD}, which takes a tuple of measurements of drefs and returns a value indicating informativeness. Thus G_{QUD} uses information from multiple dimensions to compress them into one overall measurement that addresses a contextually salient degree QUD.

In the case of (45)/(89) (see Fig. 3), maximal informativeness is distanced-based. G_{QUD} is maximal when the cardinalities of both the agent-sum and the theme-sum are maximal (see (91)). Thus the application of $\mathbf{M}_{\text{QUD-INFO}}$ selects out the maximal boy-sum and movie-sum that are involved in seeing events:

(91) $\mathbf{M}_{\text{QUD-INFO}} \stackrel{\text{def}}{=} \lambda m.\lambda g.\{h \in m(g) \mid \neg\exists h' \in m(g).$
$G_{\text{QUD}}(\langle \textsc{card}(h'(u)), \textsc{card}(h'(\nu))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \textsc{card}(h(u)), \textsc{card}(h(\nu))\rangle)\}$

Alternatively, we can consider that a sentence like (89) introduces an event dref e, satisfying the restrictions $\textsc{see}(e)$, $\textsc{boy}(\textsc{agent}(e))$, and $\textsc{movie}(\textsc{theme}(e))$. The same event dref is measured along two dimensions: agent and theme cardinality. The application of (92) selects the maximal seeing-event with the maximal agent and theme cardinality.

(92) $$\mathbf{M}_{\textsc{qud-info}} \stackrel{\text{def}}{=} \lambda m.\lambda g.\{h \in m(g) \mid \neg\exists h' \in m(g).$$
$$G_{\textsc{qud}}(\langle \textsc{card}(\textsc{agent}(h'(u))), \textsc{card}(\textsc{theme}(h'(u)))\rangle) >_{\textsc{info}}$$
$$G_{\textsc{qud}}(\langle \textsc{card}(\textsc{agent}(h(u))), \textsc{card}(\textsc{theme}(h'(u)))\rangle)\}$$

4.3 The meaning of a multi-head comparative

By combining the above analysis of comparatives and multi-measurement sentences, I show how to derive the meaning of these representative examples:

(93)
a. More dogs ate more rats than cats ate mice. (= (2))
b. Fewer people own more wealth here than elsewhere. (=(12))
c. More goods are carried faster than before. (=(30))

1. *More dogs ate more rats than cats ate mice* ((2), see Fig. 5) In this multi-head comparative, the increases along two dimensions together address a contextual degree QUD with distance-based informativeness. Thus at both the matrix and *than*-clause level, the mereologically maximal agent-sums and theme-sums that satisfy all relevant restrictions are selected. From the *than*-clause to the matrix clause, there are parallel increases along both the agent and theme cardinality, leading to an increase of informativeness in addressing the underlying degree QUD.

The derivation is shown in (94):

(94) More^{u} dogs ate more$^{\nu}$ rats than ~~*m*-many~~$^{u'}$ cats ate ~~*n*-many~~$^{\nu'}$ mice (see Fig. 5)

$$\underbrace{\mathbf{more}_{u,u'}[\mathbf{more}_{\nu,\nu'}}_{(94c)}[\underbrace{\mathbf{M}_{\textsc{qud-info}}}_{(94b)}\underbrace{[\![\text{some}^{u}\text{ dogs ate some}^{\nu}\text{ rats, some}^{u'}\text{ cats ate some}^{\nu'}\text{ mice}]\!]}_{(94a)}]]$$

a. **Introducing drefs**:

$[\![$some^{u} dogs ate some$^{\nu}$ rats, some$^{u'}$ cats ate some$^{\nu'}$ mice$]\!]$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x\\ \nu'\mapsto y'\\ u'\mapsto x'}} \;\middle|\; \textsc{rat}(y), \textsc{dog}(x), \textsc{ate}(x,y), \textsc{mouse}(y'), \textsc{cat}(x'), \textsc{ate}(x',y') \right\}$$

b. **Selecting drefs that lead to degree-QUD-based maximal informativeness**

Underlying degree QUD: $\underline{\#\text{agent} \uparrow, \#\text{theme} \uparrow}$ $\rightsquigarrow$ informativeness $\uparrow$

Thus, agents and themes with the largest cardinality are selected.

$$\mathbf{M}_{\text{QUD-INFO}} \overset{\text{def}}{=} \lambda m.\lambda g.\{h \in m(g) \mid \neg\exists h' \in m(g).$$
$$G_{\text{QUD}}(\langle \textsc{card}(h'(u)), \textsc{card}(h'(\nu))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \textsc{card}(h(u)), \textsc{card}(h(\nu))\rangle) \vee$$
$$G_{\text{QUD}}(\langle \textsc{card}(h'(u')), \textsc{card}(h'(\nu'))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \textsc{card}(h(u')), \textsc{card}(h(\nu'))\rangle)\}$$

$$\mathbf{M}_{\text{QUD-INFO}}[(94a)] \Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x\\ \nu'\mapsto y'}}_{u'\mapsto x'} \middle| \begin{array}{l} y' = \Sigma y'[\textsc{mouse}(y') \wedge \textsc{ate}(x', y')] \\ x' = \Sigma x'[\textsc{cat}(x') \wedge \textsc{ate}(x', y')] \\ y = \Sigma y[\textsc{rat}(y) \wedge \textsc{ate}(x, y)] \\ x = \Sigma x[\textsc{dog}(x) \wedge \textsc{ate}(x, y)] \end{array} \right\}$$

c. **Multi-dimensional comparisons**

$$\mathbf{more}_{u,u'} \overset{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } \textsc{card}(g(u)) \subseteq \iota I[I - \textsc{card}(g(u')) = (0, +\infty)] \\ \varnothing & \text{otherwise} \end{cases}$$

$$\mathbf{more}_{\nu,\nu'} \overset{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } \textsc{card}(g(\nu)) \subseteq \iota I[I - \textsc{card}(g(\nu')) = (0, +\infty)] \\ \varnothing & \text{otherwise} \end{cases}$$

$$\mathbf{more}_{u,u'}[\mathbf{more}_{\nu,\nu'}[(94b)]]$$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x\\ \nu'\mapsto y'}}_{u'\mapsto x'} \middle| \begin{array}{l} y' = \Sigma y'[\textsc{mouse}(y') \wedge \textsc{ate}(x', y')] \\ x' = \Sigma x'[\textsc{cat}(x') \wedge \textsc{ate}(x', y')] \\ y = \Sigma y[\textsc{rat}(y) \wedge \textsc{ate}(x, y)] \\ x = \Sigma x[\textsc{dog}(x) \wedge \textsc{ate}(x, y)] \end{array} \right\},$$

$$\text{if } \begin{array}{l} \textsc{card}(g(u)) \subseteq \iota I[I - \textsc{card}(g(u')) = (0, +\infty)] \\ \textsc{card}(g(\nu)) \subseteq \iota I[I - \textsc{card}(g(\nu')) = (0, +\infty)] \end{array}$$

In (94), drefs are first introduced, and relevant restrictions are added onto them (see (94a)). In (94b), the degree-QUD-based maximality operator picks out those drefs that lead to maximal informativeness at the *than*-clause and matrix clause respectively (see Fig. 5). In the current case, this step amounts to picking out the mereologically largest sums (with all relevant restrictions). Finally, in (94c), $\mathbf{more}_{u,u'}$ and $\mathbf{more}_{\nu,\nu'}$ work in a way similar to the cardinality tests for cumulative-reading sentences (see (89c)), checking whether there are two increases along the dimensions of agent and theme cardinality. The definition of the two tests follows the simplification shown in (85).

Alternatively, we can consider that the maximal dogs-eating-rats and cats-eating-mice events are selected and compared along agent and theme cardinality (see also (92)). The derived meaning is the same as shown in (94).

2. ***Fewer people own more wealth here than elsewhere*** **((12), see Fig. 6)** In this multi-head comparative, the decrease of owner population size and increase of owned wealth quantity together address a degree QUD with angle-based informativeness. Thus at both the matrix and *than*-clause level, the maximal drefs that lead to the maximal proportion between owned wealth quantity and owner population are selected. From the *than*-clause to the matrix clause, there is a decrease of owner population and an increase of owned wealth quantity, leading to an increase of informativeness in addressing the degree QUD (e.g., the skewness of wealth distribution).

The derivation is shown in (95):

(95) Fewer^{u} people own more$^{\nu}$ wealth here than ~~m-many$^{u'}$ people own n-much$^{\nu'}$ wealth~~ elsewhere (see Fig. 6)

$$\underbrace{\textbf{fewer}_{u,u'}[\textbf{more}_{\nu,\nu'}}_{\text{(95c)}}[\underbrace{\mathbf{M}_{\text{QUD-INFO}}}_{\text{(95b)}}\underbrace{[\![\text{some}^{u}\text{ poh own some}^{\nu}\text{ woh, some}^{u'}\text{ poe own some}^{\nu'}\text{ woe}]\!]}_{\text{(95a)}}]]$$

(I use *poh*, *woh*, *poe*, and *woe* to stand for *people-of-here*, *wealth-of-here*, *people-of-elsewhere*, and *wealth-of-elsewhere*.)

a. **Introducing drefs**:

$[\![$some^{u} poh own some$^{\nu}$ woh, some$^{u'}$ poe own some$^{\nu'}$ woe$]\!]$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x\\ \nu'\mapsto y'\\ u'\mapsto x'}} \,\middle|\, \textsc{woh}(y), \textsc{poh}(x), \textsc{own}(x,y), \textsc{woe}(y'), \textsc{poe}(x'), \textsc{own}(x',y') \right\}$$

b. **Selecting drefs that lead to degree-QUD-based maximal informativeness**

Underlying degree QUD: <u>#owners ↓, #wealth ↑</u> $\rightsquigarrow$ informativeness ↑

Thus, drefs.

$$\mathbf{M}_{\text{QUD-INFO}} \overset{\text{def}}{=} \lambda m.\lambda g.\{h \in m(g) \mid \neg\exists h' \in m(g).$$
$$G_{\text{QUD}}(\langle \mu_p(h'(u)), \mu_w(h'(\nu))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \mu_p(h(u)), \mu_w(h(\nu))\rangle) \vee$$
$$G_{\text{QUD}}(\langle \mu_p(h'(u')), \mu_w(h'(\nu'))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \mu_p(h(u')), \mu_w(h(\nu'))\rangle)\}$$

$\mathbf{M}_{\text{QUD-INFO}}[(95a)]$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x\\ \nu'\mapsto y'\\ u'\mapsto x'}} \,\middle|\, \begin{array}{l} y' = \Sigma y'[\textsc{woe}(y') \wedge \textsc{own}(x',y') \wedge \frac{\mu_w(y')}{\mu_p(x')} \text{ is maximal}] \\ x' = \Sigma x'[\textsc{poe}(x') \wedge \textsc{own}(x',y') \wedge \frac{\mu_w(y')}{\mu_p(x')} \text{ is maximal}] \\ y = \Sigma y[\textsc{woh}(y) \wedge \textsc{own}(x,y) \wedge \frac{\mu_w(y)}{\mu_p(x)} \text{ is maximal}] \\ x = \Sigma x[\textsc{poh}(x) \wedge \textsc{own}(x,y) \wedge \frac{\mu_w(y)}{\mu_p(x)} \text{ is maximal}] \end{array} \right\}$$

c. **Multi-dimensional comparisons**

$$\textbf{fewer}_{u,u'} \overset{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } \mu_p(g(u)) \subseteq \iota I[I - \mu_p(g(u')) = (-\infty, 0)] \\ \varnothing & \text{otherwise} \end{cases}$$

$$\mathbf{more}_{\nu,\nu'} \overset{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } \mu_w(g(\nu)) \subseteq \iota I[I - \mu_w(g(\nu')) = (0,+\infty)] \\ \varnothing & \text{otherwise} \end{cases}$$

$$\mathbf{fewer}_{u,u'}[\mathbf{more}_{\nu,\nu'}[(95b)]]$$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto y\\ u\mapsto x\\ \nu'\mapsto y'\\ u'\mapsto x'}} \,\middle|\, \begin{array}{l} y' = \Sigma y'[\textsc{woe}(y') \wedge \textsc{own}(x',y') \wedge \frac{\mu_w(y')}{\mu_p(x')} \text{ is maximal}] \\ x' = \Sigma x'[\textsc{poe}(x') \wedge \textsc{own}(x',y') \wedge \frac{\mu_w(y')}{\mu_p(x')} \text{ is maximal}] \\ y = \Sigma y[\textsc{woh}(y) \wedge \textsc{own}(x,y) \wedge \frac{\mu_w(y)}{\mu_p(x)} \text{ is maximal}] \\ x = \Sigma x[\textsc{poh}(x) \wedge \textsc{own}(x,y) \wedge \frac{\mu_w(y)}{\mu_p(x)} \text{ is maximal}] \end{array} \right\},$$

$$\text{if } \begin{array}{l} \mu_p(g(u)) \subseteq \iota I[I - \mu_p(g(u')) = (-\infty,0)] \\ \mu_w(g(\nu)) \subseteq \iota I[I - \mu_w(g(\nu')) = (0,+\infty)] \end{array}$$

In (95), drefs are first introduced, and relevant restrictions are added onto them (see (95a)). In (95b), the degree-QUD-based maximality operator picks out those drefs that lead to maximal informativeness at the *than*-clause and matrix clause respectively (see Fig. 6). In the current case, this step amounts to picking out the maximal wealth-owner-sum and the maximal owned-wealth such that there is (i) an owning relation between them and (ii) the ratio between their measurements is maximal. Finally, in (95c), $\mathbf{fewer}_{u,u'}$ checks whether there is a decrease along the measurement of owning population, and $\mathbf{more}_{\nu,\nu'}$ checks whether there is an increase along the measurement of their owned wealth.

As discussed earlier (see §2.2.2), just like in regular comparatives, in multi-head comparatives, the units involved in measurement can be flexible and unspecified. Thus in (95), there is no need to specify what units are involved in the measure functions μ_p (the measurement of population) or μ_w (the measurement of wealth). The percentage-based reading is available based on contextual specification of measurement units (see Krifka 1999 for a similar view in interpreting cumulative reading sentences like (50)).

3. *More goods are carried faster* (*than before*) (see (30)) Sometimes, in a multi-head comparative, the same dref is measured along different dimensions, and changes along these dimensions together address an underlying degree QUD (see also (92)).

Multi-head comparative (30) means that transported goods are measured along the dimensions of speed and quantity, and the parallel increases address an underlying QUD on the degree of transport efficiency. The derivation is shown in (96):

(96) More^u goods are carried faster$^\nu$ ~~than m-many$^{u'}$ goods were carried $n^{\nu'}$-fast~~

$$\underbrace{\mathbf{more}_{u,u'}[\mathbf{more}_{\nu,\nu'}}_{(96c)}[\underbrace{\mathbf{M}_{\text{QUD-INFO}}}_{(96b)}\underbrace{[\![\text{some}^{u} \text{ cg are } \nu\text{-fast, some}^{u'} \text{ pg were } \nu'\text{-fast}]\!]}_{(96a)}]]$$

(I use *cg* and *pg* to stand for *current goods* and *previous goods*.)

a. **Introducing drefs**:

$[\![\text{some}^{u} \text{ cg are } \nu\text{-fast, some}^{u'} \text{ pg were } \nu'\text{-fast}]\!]$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto I\\ u\mapsto X\\ \nu'\mapsto I'\\ u'\mapsto X'}} \;\middle|\; \begin{array}{l} \textsc{cg}(X), \forall x \sqsubseteq_{\text{ATOM}} X[\textsc{speed}(x) \subseteq I], \\ \textsc{pg}(X'), \forall x' \sqsubseteq_{\text{ATOM}} X'[\textsc{speed}(x') \subseteq I'] \end{array} \right\}$$

b. **Selecting drefs that lead to degree-QUD-based maximal informativeness**

Underlying degree QUD: $\underline{\text{amount-of-goods} \uparrow, \text{speed} \uparrow} \rightsquigarrow \text{informativeness} \uparrow$

Thus, mereologically maximal goods and the intervals representing their most informative speed measurement are selected.

$\mathbf{M}_{\text{QUD-INFO}} \overset{\text{def}}{=} \lambda m.\lambda g.\{h \in m(g) \mid \neg\exists h' \in m(g).$
$G_{\text{QUD}}(\langle \mu_g(h'(u)), \textsc{speed}(h'(\nu))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \mu_g(h(u)), \textsc{speed}(h(\nu))\rangle) \vee$
$G_{\text{QUD}}(\langle \mu_g(h'(u')), \textsc{speed}(h'(\nu'))\rangle) >_{\text{INFO}} G_{\text{QUD}}(\langle \mu_g(h(u')), \textsc{speed}(h(\nu'))\rangle)\}$

$$\mathbf{M}_{\text{QUD-INFO}}[(96a)] \Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto I\\ u\mapsto X\\ \nu'\mapsto I'\\ u'\mapsto X'}} \;\middle|\; \begin{array}{l} I' = \iota I'[\forall x' \sqsubseteq_{\text{ATOM}} X'[\textsc{speed}(x') \subseteq I']] \\ X' = \Sigma X'[\textsc{pg}(X')] \\ I = \iota I[\forall x \sqsubseteq_{\text{ATOM}} X[\textsc{speed}(x) \subseteq I]] \\ X = \Sigma X[\textsc{cg}(X)] \end{array} \right\}$$

c. **Multi-dimensional comparisons**

$$\mathbf{more}_{u,u'} \overset{\text{def}}{=} \lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } \mu_g(g(u)) \subseteq \iota I[I - \mu_g(g(u')) = (0, +\infty)] \\ \varnothing & \text{otherwise} \end{cases}$$

$$\mathbf{more}_{\nu,\nu'} \overset{\text{def}}{=}$$

$$\lambda m_{\langle s,st\rangle}.\lambda g_s.\begin{cases} m(g) & \text{if } \textsc{speed}(g(\nu)) \subseteq \iota I[I - \textsc{speed}(g(\nu')) = (0, +\infty)] \\ \varnothing & \text{otherwise} \end{cases}$$

$\mathbf{more}_{u,u'}[\mathbf{more}_{\nu,\nu'}[(96b)]]$

$$\Leftrightarrow \lambda g.\left\{ g^{\substack{\nu\mapsto I\\ u\mapsto X\\ \nu'\mapsto I'\\ u'\mapsto X'}} \;\middle|\; \begin{array}{l} I' = \iota I'[\forall x' \sqsubseteq_{\text{ATOM}} X'[\textsc{speed}(x') \subseteq I']] \\ X' = \Sigma X'[\textsc{pg}(X')] \\ I = \iota I[\forall x \sqsubseteq_{\text{ATOM}} X[\textsc{speed}(x) \subseteq I]] \\ x = \Sigma X[\textsc{cg}(X)] \end{array} \right\},$$

$$\text{if } \begin{array}{l} \mu_g(g(u)) \subseteq \iota I[I - \mu_g(g(u')) = (0, +\infty)] \\ \textsc{speed}(g(\nu)) \subseteq \iota I[I - \textsc{speed}(g(\nu')) = (0, +\infty)] \end{array}$$

In (96), drefs (that are entities and intervals) are first introduced, and relevant restrictions are added onto them (see (96a)). In (96b), the degree-QUD-based maximality operator picks out those drefs that lead to maximal informativeness at the *than*-clause and matrix clause respectively. In the current case, this step amounts to picking out the mereologically maximal sums of goods and the most informative intervals that represents their speed (see (78) for the definition of 'most informative interval', repeated here in (97)). Finally, in (96c), $\mathbf{more}_{u,u'}$ and $\mathbf{more}_{\nu,\nu'}$ check whether there are parallel increases along the dimensions of goods quantity and transport speed.

(97) Selecting the most informative interval from a set of intervals p s.t.
$\exists I[p(I) \wedge \forall I'[[p(I') \wedge I' \neq I] \to I \subset I']]$
When defined, The most informative interval is
$\iota I[p(I) \wedge \forall I'[[p(I') \wedge I' \neq I] \to I \subset I']]$ (= (78))

This example analyzed in (96) is slightly different from the cases analyzed in (94) and (95) in that there is no mutual restriction between the entity dref and the interval dref, at the matrix or *than*-clause level. As shown in (96b), drefs I and I' are not involved in restricting drefs X and X'. In this sense, most informative entity drefs X and X' are selected first, before most informative interval drefs I and I' are selected.

This difference between (96) and (94)/(95) is due to the dimensions along which measurements are taken. In cumulative-reading sentences (see (94)/(95)), multi-measurements are along two dimensions of quantity, leading to mutual restriction along these two dimensions and the selection of mereologically restricted (not absolute) maximal drefs. On the other hand, for (96), only one measurement is along the dimension of quantity, so that there is no mutual restriction and the mereologically absolutely largest sums of previous goods and current goods are selected.

Nevertheless, the use of informativeness-based maximality operator $\mathbf{M}_{\text{QUD-INFO}}$ provides a unified analysis that uses multi-dimensional measurements to compute informativeness.

As mentioned earlier, there is an alternative way to derive the meaning of multi-head comparatives that are based on cumulative-reading sentences (e.g., the cases in (94)/(95)): With the use of event drefs (see (92)), mutual restriction between drefs can be removed, and the maximal event drefs at the matrix and *than*-clause level are selected and compared along different dimensions.

In general, multi-head comparatives are not necessarily based on cumulative-reading sentences or the measurement along two dimensions of quantity (see also the discussion in §2.2.3):

(98) More people spent more of their time focusing on what they're thankful for. (= (25))

a. **Dimensions**: human quantity and the length of time spent per person

b. $\rightsquigarrow$ Selecting and comparing (i) the maximal human-sum that spent time focusing on what they're thankful for, ΣX, and (ii) the most informative interval I s.t. $\forall x \sqsubseteq_{\text{ATOM}} X[\mu_{\text{TIME}}(x) \subseteq I]$

(99) Aldi's pork selection was a little more varied with better value than what I saw here. (= (31))

a. **Dimensions**: how varied pork selection is and how valuable it is

b. $\rightsquigarrow$ Aldi's and what I saw here are compared along these two dimensions

(100) Newer generations of microchips contain more electronic switches on a smaller surface. (= (39b))

a. **Dimensions**: how many electronic switches a microchip has and the surface of a microchip

b. $\rightsquigarrow$ Older and newer microchips are compared along these two dimensions.

(98) is similar to (96) in involving only one dimension of quantity. Thus, there is no mutual restriction along the two dimensions. The mereologically largest sums of relevant humans (for the contextually accommodated *than*-clause and the matrix clause) are selected and compared along the two dimensions in (98a), thus deriving the meaning of a multi-head comparative with distributive reading (see §2.2.2).

In interpreting (99) and (100), there is no dimension of quantity, and thus mereologically maximal sums are not necessarily involved at all. In particular, (100) can be simply used to make a generic claim about older and newer generations of microchips and there is no need to select the maximal sums of microchips.

5 Discussion

In the above proposed analysis, the current work inherits insights from the existing literature, especially von Stechow (1984)'s pioneering work on (multi-head) comparatives and Brasoveanu (2013)'s formal analysis of cumulative-reading sentences. However, the current analysis takes a new perspective on the notions of comparison and informativeness, paving the ground for analyzing more empirical data.

5.1 Comparison: additivity vs. inequality

According to von Stechow (1984), the semantics of comparison as expressed by comparatives is essentially built on an inequality relation between two values, one of which is considered the comparison standard and contributed by a *than*-clause/phrase.

Thus given that a multi-head comparative expresses multiple comparisons, but at most one *than*-clause/phrase is explicitly uttered, von Stechow (1984) proposes that there is reconstruction (see (101)). Eventually, the derivation of the semantics of a multi-head comparative is based on multiple *than*-clauses.

(101) **Under von Stechow (1984)'s proposal of reconstruction** (p.43: (144)):
More dogs [than cats ate mice] ate more rats than cats ate mice. (= (2))

Under the current analysis, (i) comparison is actually the same phenomenon as additivity, extending an existing, salient base value to higher informativeness, and (ii) comparative morphemes like *more*/*-er* and *fewer*/*less* are similar to additive particle *another* and denote a positive or negative increase. Thus comparative morphemes *more*/*-er* do not work like '>' (which is of type $\langle d, \langle dt \rangle\rangle$ and requires two arguments), and no reconstruction of another *than*-clause is needed.

Technically, the current additivity-based perspective on comparison does not need to explain how reconstruction happens syntactically or why the reconstructed *than*-clause can never be overtly uttered (see (101)). Empirically, the semantics of a multi-head comparative is not considered two inequalities, but rather increases or decreases along two dimensions, making it conceptually easier to entertain the interplay between increases / decreases and the contextual degree QUD behind this interplay.

Under the current analysis in §4.3, regular, single-dimensional comparatives can be considered a special case of multi-head comparatives: degree-QUD-based maximal informativeness (see (90)) is computed from one-dimensional measurement.

Besides, given that *-er*/*more* denotes an increase (i.e., $(0, +\infty)$), it can further be restricted by expressions like *slightly* and *much*, just like *-er*/*more* can be restricted by a numerical differential in a regular comparative (e.g., *2 inches* in (82)/(83)/(84)).

Thus a multi-head comparative like (102) resolves its underlying degree QUD with two increases that are disproportionate (*slightly* vs. *much*). (See also *a little more varied* in the naturally occurring example (31)/(99).)

(102) Slightly better products^u charge much higher price than before.

$\rightsquigarrow$ products are measured and compared along the degrees of quality and price

a. Along the dimension of quality, the default difference $(0, +\infty)$ is restricted by $(0, d]$ (here d is the threshold of being small for a difference of quality)

b. Along the dimension of price, the default difference $(0, +\infty)$ is restricted by $[d', +\infty)$ (here d' is the threshold of being large for a difference of price)

5.2 Degree QUD-based informativeness maximality

Within Brasoveanu (2013)'s analysis of cumulative-reading sentences, mereologically maximal drefs that satisfy all relevant restrictions are selected for checking cardinality (see §3.1). The current analysis follows Krifka (1999) and Zhang (2023) and considers mereology-based maximality a special case of informativeness-based maximality: cumulative-reading sentences and multi-head comparatives integrate information along multiple dimensions to address an underlying degree QUD.

As a consequence, under the current proposal, a degree-QUD-based maximality operator is actually similar to focus sensitive particles like *even*.

In (103a), the uttered modified numerals invoke focus alternatives in addressing the underlying degree QUD and resolve the QUD with maximal informativeness (see also Krifka 1999; cf. Brasoveanu 2013). Similarly, in (103b), the associate of *even* (i.e., *3*) also invokes focus alternatives in addressing the underlying degree QUD and resolves the QUD with maximal informativeness. (103b) is typically not used to mean that it's unlikely for Mary or anyone else to have 3 kids. Rather, (103b) means that having 3 kids represents reaching the highest level in resolving a relevant degree QUD.

(103) a. At most 3% of the population own at least 70% of the land. (= (50))
Degree QUD: *how skewed wealth distribution is* (see Fig. 4)

b. Mary has even $\underline{3}_F$ kids. (see Greenberg 2018, Zhang 2022)
Degree QUD: *how heavy Mary's childcare burden is*

Thus the current analysis predicts that *even* can be used in a cumulative-reading sentence or a multi-head comparative to address the same underlying degree QUD. This prediction is borne out. As illustrated in (104), both the use of *even* and the use of numerals or changes express the skewness of wealth distribution:[8]

[8] I thank Sigrid Beck for pointing out this to me.

(104) a. At most 3% of the population owns even more than 70% of the land.
b. Fewer and fewer people own even more and more wealth.

5.3 A related phenomenon: Comparative correlatives

Under the additivity-based perspective, analyzing the semantics of *more*/*-er* and *less* as an increase / decrease also immediately provides an analysis of comparative correlatives.

Comparative correlatives are similar to multi-head comparatives in containing two comparative expressions, but comparative correlatives never contain a *than*-clause.

With an uttered or accommodated *than*-clause, a multi-head comparative addresses an increase of degree-QUD-based informativeness (from the *than*-clause to the matrix level). Without any *than*-clauses, comparative correlatives (see (105) and (106)) indicate how the change along one dimension determines the change along the other dimension:

(105) The more I know about my dog, the better I like her. (Zhang and Ling 2021)
$\rightsquigarrow$ the increase along the scale of familiarity determines the increase along the scale of fondness.

(106) The more you buy, the lower unit price you pay.
$\rightsquigarrow$ the increase along the scale of purchase amount determines the decrease along the scale of unit price.

6 Conclusion

Based on the understanding of comparatives and cumulative-reading sentences, I have discussed and analyzed the semantics and pragmatics of multi-head comparatives. A multi-head comparative expresses increases / decreases along multiple dimensions, and the interplay between changes along multiple dimensions addresses a single underlying issue. I have developed and explored the notion of 'degree-QUD-based informativeness' to characterize this single underlying issue.

Multi-head comparatives shed light on how human cognition deals with complex mathematical operations that involve measurement and comparison along multiple dimensions. In this regard, the current work also echoes existing research on intuitive mathematical computations encoded in language expressions (e.g., Zhang and Ling 2021 on interval subtraction, Coppock 2022 on division).

Acknowledgments

For comments and discussions on data and ideas, I thank Duk-Ho An, Sigrid Beck, Seth Cable, Zhuang Chen, Yael Greenberg, Heidi Harley, Jeremy Kuhn, Takeo Kurafuji, Chungmin Lee, Daiki Matsuoka, Toshiko Oda, Haihua Pan, Osamu Sawada, Guillaume Thomas, Carla Umbach, Alexander Williams, Yusuke Yagi, Florence Zhang, Eytan Zweig, and the audience at the symposium on 'Comparative constructions in English and other languages' at *the 41st Annual Conference of the English Linguistics Society of Japan* (University of Tokyo, 2023) and *the 35th European Summer School in Logic, Language, and Information* (*ESSLLI 2024*, KU Leuven, Belgium, July-August, 2024). I specially thank Toshiko Oda for organizing the above symposium and encouraging me to work on this topic. I gratefully acknowledge financial support from NYU Shanghai and Shanghai Oriental Talent Program (PI: LZ). I also thank my parents and my friends (especially Jia Ling, Jinshuai Cao, Li Bai, and Quan Gan). All errors are mine.

References

Beaver, David I., and Brady Z. Clark. 2009. *Sense and sensitivity: How focus determines meaning*, volume 12. John Wiley & Sons. https://onlinelibrary.wiley.com/doi/book/10.1002/9781444304176.

Beck, Sigrid. 2011. Comparative constructions. In *Semantics*, ed. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, volume 2, 1341–1390. Berlin, Boston: de Gruyter. https://doi.org/10.1515/9783110255072.1341.

Brasoveanu, Adrian. 2013. Modified numerals as post-suppositions. *Journal of Semantics* 30:155–209. https://doi.org/10.1093/jos/ffs003.

Bumford, Dylan. 2017. Split-scope definites: Relative superlatives and Haddock descriptions. *Linguistics and Philosophy* 40:549–593. https://doi.org/10.1007/s10988-017-9210-2.

Büring, Daniel. 2003. On d-trees, beans, and b-accents. *Linguistics and Philosophy* 26:511–545. https://doi.org/10.1023/A:1025887707652.

Charlow, Simon. 2017. Post-suppositions and semantic theory. https://ling.auf.net/lingbuzz/003243, *Journal of semantics* (in press).

Chomsky, Noam. 1981/1993. *Lectures on government and binding*. Foris, Dordrecht; Mouton de Gruyter. https://doi.org/10.1515/9783110884166.

Coppock, Elizabeth. 2022. Division vs. distributivity: Is *per* just like *each*? In *Semantics and Linguistic Theory*, volume 32, 384–403. https://doi.org/10.3765/salt.v1i0.5335.

Davies, Mark. 2008-. *The Corpus of Contemporary American English (COCA)*. https://www.english-corpora.org/coca/.

Dayal, Veneeta. 1996. *Locality in WH quantification: Questions and relative clauses in Hindi*. Springer Science & Business Media. https://doi.org/10.1007/978-94-011-4808-5.

Fintel, Kai Von, Danny Fox, and Sabine Iatridou. 2014. Definiteness as maximal informativeness. In *The art and craft of semantics: A festschrift for Irene Heim*, ed. Luka Crnič and Uli Sauerland, volume 2, 175 – 179. Cambridge, MA. https://semanticsarchive.net/Archive/jZiNmM4N/.

Greenberg, Yael. 2010. Additivity in the domain of eventualities (or: Oliver twist's *more*). In *Proceedings of Sinn und Bedeutung*, volume 14, 151–167. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/465.

Greenberg, Yael. 2018. A revised, gradability-based semantics for even. *Natural Language Semantics* 26:51–83. https://doi.org/10.1007/s11050-017-9140-0.

Groenendijk, Jeroen, and Martin Stokhof. 1990. Two theories of dynamic semantics. In *European workshop on logics in artificial intelligence*, 55–64. Springer. https://eprints.illc.uva.nl/id/eprint/1160/1/LP-1990-14.text.pdf.

Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 39–100. https://www.jstor.org/stable/25001418.

Haddock, Nicholas J. 1987. Incremental interpretation and combinatory categorial grammar. In *The 10th International Joint Conference on Artificial Intelligence*, volume 2, 661–663. Morgan Kaufmann Publishers Inc. https://www.ijcai.org/Proceedings/87-2/Papers/012.pdf.

Hausser, Roland. 1983. The syntax and semantics of English mood. In *Questions and answers*, ed. F. Kiefer, 97–158. Dordrecht: Springer.

Hausser, Roland, and Dietmar Zaefferer. 1978. Questions and answers in a context-dependent Montague grammar. In *Formal semantics and pragmatics for natural languages*, 339–358. Springer.

Heim, Irene. 1985. Notes on comparatives and related matters. https://semanticsarchive.net/Archive/zc0ZjY0M/Comparatives%2085.pdf, unpublished ms., University of Texas, Austin.

Hendriks, Petra. 1994. Multiple head comparison and infinite regress. In *Proceedings of the Western Conference On Linguistics (WECOL 92)*, ed. Joel Ashmore Nevis and Vida Samiian, volume 5, 117–131. https://www.researchgate.net/publication/2814411_Multiple_Head_Comparison_and_Infinite_Regress.

Hendriks, Petra, and Helen De Hoop. 2001. Optimality theoretic semantics. *Linguistics and Philosophy* 24:1–32. http://www.jstor.org/stable/25001802.

Kennedy, Christopher. 1999. *Projecting the adjective*. New York: Routledge. https://doi.org/10.4324/9780203055458.

Krifka, Manfred. 1999. At least some determiners aren't determiners. In *The semantics/pragmatics interface from different points of view*, ed. K. Turner, 257–291. Elsevier. https://semantics.uchicago.edu/kennedy/classes/w14/implicature/readings/krifka99.pdf.

Marques, Rui. 2005. Multi-headed comparatives in Portuguese. *Journal of Portuguese Linguistics* 5 (1):5–35. https://doi.org/10.5334/jpl.153.

Meier, Cécile. 2001. Multihead comparatives and result clause constructions with 'split anteced- ents'. In *Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow*, ed. Caroline Féry and Wolfgang Sternefeld, 348–371. Berlin: Akademie-Verlag. https://doi.org/10.1515/9783050080116.

Moore, Ramon E. 1979. *Methods and Applications of Interval Analysis*. SIAM. https://epubs.siam.org/doi/book/10.1137/1.9781611970906.

Oda, Toshiko. 2008a. Degree Constructions in Japanese. Doctoral Dissertation, University of Conneticut. https://digitalcommons.lib.uconn.edu/dissertations/AAI3345201/.

Oda, Toshiko. 2008b. Multihead Comparatives in Japanese. *The Journal of Humanities and Natural Sciences* 126:3–31. http://hdl.handle.net/11150/495.

Roberts, Craige. 1996/2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and pragmatics* 5:6:1–69. https://doi.org/10.3765/sp.5.6.

Schwarzschild, Roger. 2008. The semantics of comparatives and other degree constructions. *Language and Linguistics Compass* 2:308–331. https://doi.org/10.1111/j.1749-818X.2007.00049.x.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3:1–77. https://doi.org/10.1093/jos/3.1-2.1.

Thomas, Guillaume. 2010. Incremental *more*. In *Semantics and Linguistic Theory*, volume 20, 233–250. https://doi.org/10.3765/salt.v20i0.2558.

Thomas, Guillaume. 2011. Another additive particle. In *Semantics and Linguistic Theory*, volume 21, 634–651. https://doi.org/10.3765/salt.v21i0.2589.

Zeevat, Henk. 2004. Particles: Presupposition triggers, context markers or speech act markers. In *Optimality theory and pragmatics*, 91–111. Springer. https://doi.org/10.1057/9780230501409_5.

Zeevat, Henk, and Katja Jasinskaja. 2007. *And* as an additive particle. In *Language, representation and reasoning. Memorial volume to Isabel Gómez Txurruka*, 315–340. University of the Basque Country Press. https://idsl1.phil-fak.uni-koeln.de/sites/IDSLI/dozentenseiten/Jasinskaja/zeevat_jasinskaja_2007.pdf.

Zhang, Linmin. 2020. Comparatives bring a degree-based NPI licenser. In *Monotonicity in Logic and Language: Proceedings of TLLM 2020*, ed. D. Deng, F. Liu, M. Liu, and D. Westerståhl, 217–238. Springer. https://doi.org/10.1007/978-3-662-62843-0_12.

Zhang, Linmin. 2022. The presupposition of *even*. In *Proceedings of Semantics and Linguistic Theory*, volume 32, TBA. https://ling.auf.net/lingbuzz/006760.

Zhang, Linmin. 2023. Cumulative reading, qud, and maximal informativeness. In *Logic & Engineering of Natural Language Semantics (LENLS)*, ed. D. Bekki et al., volume 19, 1–17. https://doi.org/10.1007/978-3-031-43977-3_1.

Zhang, Linmin, and Jia Ling. 2021. The semantics of comparatives: A difference-based approach. *Journal of Semantics* 38:249–303. https://doi.org/10.1093/jos/ffab003.

Zhang, Linmin, and Florence Yukun Zhang. 2024. Comparative morphemes are additive particles: English *-er*/*more* vs. Chinese *gèng*. https://ling.auf.net/lingbuzz/008122, *Generative Perspectives on Degrees*.