The mereological structure of distributivity: a case study of binominal *each*

Jess H.-K. Law

University of California, Santa Cruz

Abstract

Binominal *each* is known to exhibit selectional requirements on the noun phrase that immediately precedes it. The goal of this paper is to reduce these selectional requirements to a single requirement of monotonic growth of measurement in relation to the 'size' of distributivity. More concretely, it is argued that binominal *each* imposes a constraint on the functional dependencies arising from distributive quantification, requiring that the measurement of its host grows monotonically with the number of values being distributively quantified.

To make constraints on dependencies formally explicit, I devise a version of dynamic plural logic with features from van den Berg (1996) and Brasoveanu (2008, 2013) to semantically represent dependencies arising from evaluating distributive quantification. The use of a dynamic logic, coupled with a delayed evaluation mechanism in terms of higher order meaning (Cresti 1995, de Swart 2000, Charlow, to appear), allows the constraint to act as an output context constraint on distributive quantification, which mirrors the use of output constraints pioneered by Farkas (1997, 2002b) and and further developed in Brasoveanu (2013), Henderson (2014) and Kuhn (2017).

1 Introduction

Each is a distributive marker in English that can occur in a few grammatical positions. It may appear as a determiner, as shown in (1), as a pre-verbal adverb, as shown in (2), or as a post-nominal element, as shown in (3). When appearing in the last position, it is often referred to as 'binominal *each*', following the terminology of Safir and Stowell (1988). [1]

(1)   Each girl saw two movies.   Determiner

(2)   The girls each saw two movies.   Adverbial

(3)   The girls saw two movies *each*.   Binominal

Many studies have recognized the distributive function of *each* in all of the three positions, which is responsible for making sure that seeing of two movies is true of every single girl being considered. They have also noted, however, that binominal *each* seems to communicate

---

[1]Other names have been given to binominal *each*, including 'anti-quantifier' (Choe 1987), 'shifted *each*' (Postal 1974), and 'adnominal *each*' (Champollion 2016). The name 'binominal *each* is adopted here to highlight the intuition, due to Safir and Stowell (1988), that this use of *each* makes reference to two noun phrases. As is argued in section 4, binominal *each* makes reference to two noun phrases via discourse anaphora.

something more than just distributivity, in the form of morpho-syntactic and interpretive requirements on its host, i.e., the noun phrase immediately preceding it.[2] These requirements, to be introduced in more detail in section 2, are briefly summarized below:

The variation requirement: binominal *each*, like markers of dependent indefinites, requires its host to co-vary with a distributive quantifier (Cable 2014, Champollion 2015, Kuhn 2017). Because of this requirement, it is infelicitous to identify two movies that every girl saw in (4). This requirement is subject to more discussion in section 2.1.

(4)   The girls saw two movies each, #namely, *Inception* and *The Godfather*.

The counting quantifier requirement: the host must have a counting component, which may be a numeral like *two* but not a quantificational determiner like *some*, as exemplified by (5) (Safir and Stowell 1988, Sutton 1993, Szabolcsi 2010). This requirement is taken up further in section 2.2.

(5)   The girls saw two/*some movies each.

The monotonic measurement requirement: the host must be associated with monotonic measurement (Zhang 2013), such as the measurement of *volume* but not the measurement of *temperature*, as exemplified by (6). This requirement is discussed further in section 2.3.

(6)   The drinks are 6 ounces/*60 degrees each.

These requirements are unique to binominal *each* and are not found with *each* in other positions. The goal of this paper is to provide a single requirement, called a **monotonicity constraint**, that subsumes these three sets of seemingly unconnected empirical observations. To simply put, for (3) this constraint requires a monotonic mapping from the number of girls to the measurement of the movies they saw. A mapping is taken to be monotonic here when it is *non-constant* and *non-decreasing*. Using (3) as a concrete example, the monotonicity constraint requires that the number of movies does not remain constant or decrease when more girls are considered.

A monotonicity constraint of this kind has an intimate connection with distributive quantification: distributive quantification can establish a **linking relation** between the values in its restriction (e.g., the girls in (3)) and the values in its scope (e.g., the movies seen in (3)). The linking relation makes it possible to talk about monotonic mappings between values in the restriction and values in the scope of distributive quantification. The same type of linking relation has been independently motivated by a host of previous studies to account for quantificational subordination, dependent indefinites, and adjectives with a sentence-internal reading (Krifka 1996a, van den Berg 1996, Nouwen 2003, Solomon 2011, Brasoveanu 2008, 2011, Henderson 2014, Kuhn 2017). I follow previous studies and refer to this kind of linking a functional dependency or a **dependency** for short. The formal, explicit construction of a dependency is done

---

[2]Following Safir and Stowell (1988), I assume that binominal *each* is base-generated in a post-nominal position and forms a constituent with its host. The motivation for this particular constituency comes from *wh*-movement and passivization:

(i)   How many girls each did the men see?   (Safir and Stowell 1988:ex.5a)

(ii)   One girl each was seen by the men.   (Safir and Stowell 1988:ex.5b)

Alternatives to the base-generation analysis have also been explored. For example, Stowell (2013) proposes an analysis that derives binominal *each* from adverbial *each*.

with a version of dynamic plural logic in section 4, Before that, I speak of a dependency arising from distributive quantification only informally, using visualizations like Figure 1 to indicate what a dependency introduced by distributive quantification looks like.
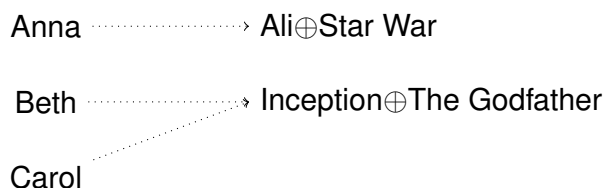

Figure 1: A sample dependency (i.e., linking relation) established by interpreting *The girls saw two movies each* in a context in which Anna saw Ali and Star War, but Beth and Carol both saw Inception and The Godfather.

Although studies abound on dependencies established by dynamically interpreting distributive quantification, a property about these dependencies remain unexplored, namely, that dependencies have a nontrivial **mereological structure**. For concreteness, the mereological structure of the dependency in Figure 1 is visualized in Figure 2. With this kind of mereological structure, we can speak of not only the girls and the movies they saw, but also how the sizes of the two pluralities relate. The mereological nature of dependencies has not played a role in previous discussions of dependencies but is crucial for formulating the monotonicity-based account of binominal *each* in this paper.

The monotonicity constraint of binominal *each* can be seen as a special way of using dependencies arising from distributive quantification—it is a constraint on **how the sizes of two mereological structures relate**. In this sense, although the semantics proposed for binominal *each* aligns it, in broad terms, with the semantics for markers of dependent indefinites devised in Henderson (2014) and Kuhn (2017), it crucially differs from them in the need to access the mereological structure of a dependency. If the current approach is on the right track, it lends support to treating distributivity as a form of plurality with a mereological structure, as advocated in the tradition of Dynamic Plural Logic (van den Berg 1996, Nouwen 2003, Brasoveanu 2008), and event-based distributivity (Schein 1993, Lasersohn 1995, Champollion 2010). After all, having a nontrivial mereological structure is precisely what plural structures are known for.

To build functional dependencies with a mereological structure, I sketch a variant of Dynamic Plural Logic that shares important features with both van den Berg (1996) and Brasoveanu (2008). Binominal *each* is able to compositionally access these dependencies after taking split
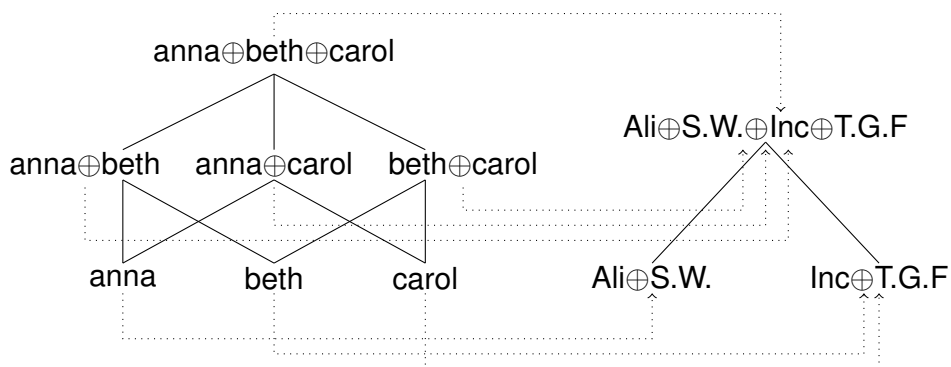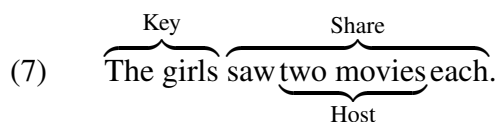

Figure 2: A dependency arising from distributive quantification has a mereological structure.

scope (together with its host) over distributivity with the help of a **higher order meaning** (Cresti 1995, de Swart 2000, Charlow, to appear). The use of higher order meaning for accessing dependencies compositionally is another feature that distinguishes the present study from previous studies on dependent indefinites (e.g., Henderson 2014, Kuhn 2017).

This paper proceeds as follows. Section 2 takes up the three selectional requirements binominal *each* imposes on its host. Section 3 offers the informal generalization that binominal *each* imposes a monotonicity constraint on the mereological structure of the dependencies arising from distributive quantification. Readers only interested in the monotonicity generalization and not how it is formally and compositionally implemented only needs to read up to the end of section 3. Section 4 provides a formal, compositional implementation of the monotonicity constraint in a dynamic plural logic. Section 5 provides a comparison of my analysis of binominal *each* with the analyses put forward in previous studies. Sections 6 and 7 discuss the nature of output-context constraints as well as potential challenges to the monotonicity generalization. Section 8 concludes.

2 The selectional requirements of binominal *each* on its host

The classical analysis of binominal *each* is that it contributes a distributive operator (Link 1987, Safir and Stowell 1988, Zimmermann 2002, Dotlačil 2012), just like its close correlate adverbial *each* (as analyzed in Link 1987, Roberts 1987, Schwarzschild 1996, among many others). However, many studies have observed additional properties of binominal *each* that do not directly follow from it being a distributive operator (e.g., Safir and Stowell 1988, Sutton 1993, Zimmermann 2002, Stowell 2013, Cable 2014, Zhang 2013). This section is devoted to familiarizing the reader with these properties. To facilitate the discussion, let me introduce some terminology for referring to different parts of a sentence with binominal *each* using the example in (7).

(7)　　　The girls $\overbrace{\text{saw}}^{\text{Key}}$ $\underbrace{\text{two movies}}_{\text{Host}}$ $\overbrace{\text{each.}}^{\text{Share}}$

The noun phrase that immediately precedes binominal *each* is called the *host* of binominal *each*. Following the terminology in previous studies on distributivity (Choe 1987, Zimmermann 2002, Champollion 2017), the noun phrase being distributively quantified, typically a plural expression occupying the subject position, is called the (*Distributive*) *Key*. The entire predicate following the Key is called the (*Distributed*) *Share*.

2.1 The variation requirement

Safir and Stowell (1988) is the earliest study, as far as I know, to notice the variation requirement of binominal *each*. They observe that in a sentence like (8), there is a strong preference that the girls did not all see the same two movies. In fact, if one tries to add a continuation clause to identify two particular movies, as in done (9), the result becomes quite odd.

(8)　　　The girls saw two movies each.

(9)　　　*The girls saw two movies each, namely *Avatar* and *Ice Age*.

Safir and Stowell (1988) treat binominal *each* as a polyadic distributive operator that quantifies over sets provided by two nominals (hence the name 'binominal'.) According to them, the

quantification in (8) yields a *one-to-one correspondence* between girls and movies, such that each girl saw a different set of two movies.[3]

Moltmann (1991) points out that the one-to-one correspondence condition is too strong. This is because a sentence like (8) is judged true not only when there is a one-to-one correspondence between the girls and the movies (or a total variation of the movies), such as the scenario shown in Figure 3 (left), but also when there is only partial variation of the movies, such as the scenario shown in Figure 3 (right).[4]

Anna ⟶ Aliens⊕Star War   Anna ⟶ Aliens⊕Star War

Beth ⟶ Inception⊕The Godfather   Beth ⟶ Inception⊕The Godfather
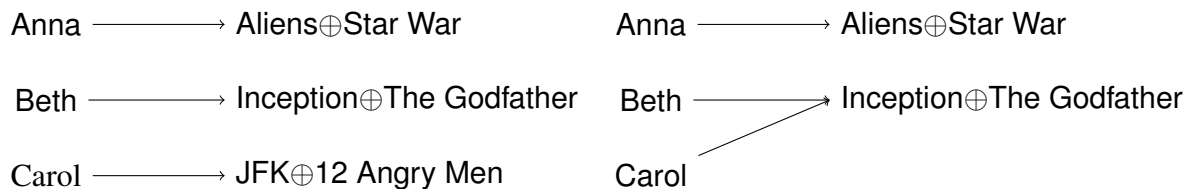
Carol ⟶ JFK⊕12 Angry Men   Carol

Figure 3: Both total and partial variation are acceptable for binominal *each*.

Moltmann suggests weakening the variation requirement to a condition of distinct discourse referents (d-refs), noting that distinct d-refs do not necessarily have distinct values. A problem with Moltmann's characterization is that since distinct d-refs do not necessarily require distinct values, they fail to model the variation requirement altogether. Recent studies settled on characterizing the variation requirement as a non-constant correspondence (Cable 2014, Champollion 2015, Kuhn 2017), an insight borrowed from work on dependent indefinites.

Generally speaking, a dependent indefinite is an indefinite noun phrase with a morphological marker that induces a distributive interpretation of the sentence.[5] In addition, the morphological marker bears an additional component requiring the indefinite to contribute a witness that co-varies with the Key.[6] The following sentence from Hungarian (cited from Farkas 1997) illustrates a distributed numeral marked by numeral reduplication:

(10)   Minden gyerek olvasott egy-egy konyvet.
        every   child   read   a-a   book
        'The children read one book each.'   Hungarian
    a.   Distributivity inference: Each child read a book.
    b.   Variation inference: More than one book was read.

Couched in various frameworks, Farkas (1997, 2002b,a), Balusu (2005) and Henderson (2014) have proposed that the variation requirement follows from a special plurality condition rooted in dependent indefinites. The special plurality condition requires a non-constant correspondence

---

[3]Safir and Stowell (1988) credits Jim Higginbotham (p.c.) for suggesting the one-to-one correspondence condition.

[4]More formally, Figure 3 (left) represents a bijective function while Figure 3 (right) represents a non-injective surjective function.

[5]Dependent indefinites are known by various names. They are called 'dependent indefinites' in Farkas (1997, 2002b,a), Henderson (2014), and Kuhn (2017) and 'distributive numerals' in Cable (2014). In languages that use reduplication to mark dependent indefinites, they are commonly referred to as 'reduplicated numerals' (Gil 1988, Balusu 2005). Since dependent indefinites signal distributivity without being close to the distributivity key, they have also been characterized as exhibiting 'distance distributivity' (Zimmermann 2002, Cable 2014).

[6]I use 'co-vary' in the same way it is used in Beghelli et al. (1997), i.e., a variables y co-varies with a variable x iff not all values of x are associated with the same value of y. This type of co-variation is also more formally referred to as 'variable dependence' in this paper. The term 'non-constant correspondence' can be seen as an informal characterization of the same concept.

between values introduced by a dependent indefinite and values being distributively quantified.[7] I review, in section 5.1, how the variation requirement of dependent indefinites is treated in the plurality approach, using Henderson (2014) as an example. It suffices at this point to know that attempts to extend the plurality treatment of dependent indefinites to binominal *each*, such as Champollion (2015) and Kuhn (2017), have been largely successful in modeling the variation requirement of binnominal *each*.[8]

However, despite the empirical adequacy in handling the variation requirement, the plurality approach fails to provide insight into two other puzzling requirements of binominal *each*, namely, the counting quantifier requirement and the monotonic measurement requirements. These requirements are discussed in the next two subsections.

2.2 The counting quantifier requirement

It is generally agreed that binominal *each* forms a constituent with its host (e.g., Burzio 1986, Safir and Stowell 1988). In addition, studies have shown that binominal *each* seem to select some forms of indefinites as its host (e.g., Safir and Stowell 1988, Zimmermann 2002, Stowell 2013). The most precise description, as far as I know, comes from Sutton (1993), who concludes that only counting quantifiers, i.e., noun phrases with (modified) numerals or vague quantity words like *many*, *a few* or *several*, can host binominal *each* (see also Szabolcsi 2010).[9] All other noun phrases are rejected as a host of binominal *each*. The contrast between counting quantifiers and other quantifiers is illustrated in (11) and (12).[10]

---

[7]The plurality requirement is weaker than a one-to-one correspondence. However, Henderson (2014:fn.15) notes that although the plurality requirement seems to be truth-conditionally adequate, native speakers of Kaqchikel have a preference for full covariation, i.e., a one-to-one correspondence. A similar preference seems to also hold for binominal *each* (Simon Charlow, p.c.).

[8]A few other strategies have been explored to model the variation requirement. In Choe (1987), the variation requirement is used to signal the obligatory narrow scope of the host of binominal *each*. However, it must be made clear that while narrow scope may give rise to co-variation, it does not guarantee it. For this reason, the variation requirement cannot be simply restated as a narrow-scope requirement, contrary to the claim in Choe (1987).

Another possibility that has been considered is to generate the variation requirement as an implicature. For concreteness, the wide scope indefinite interpretation entails the narrow-scope indefinite interpretation. By using binominal *each* to explicitly signal the narrow-scope indefinite interpretation, one indicates that the wide scope indefinite interpretation is false, hence triggering a variation implicature. This possibility is briefly considered in Henderson (2014) and discussed in more detail in Kuhn (2015: Ch.3.6). However, both Henderson (2014) and Kuhn (2017) reject a scalar implicature account for the variation requirement. Henderson's main objection is that scalar implicatures should be cancellable when the context fails to support it. However, when the variation requirement of dependent indefinites and binominal *each* fails, ungrammaticality is observed. Kuhn cautions using cancellability to diagnose scalar implicatures, as more recent studies have identified a host of grammaticalized, obligatory scalar implicatures (see Chierchia (2006), Chierchia et al. 2011, Fox 2007, a.o.). Using data from American Sign Language, Kuhn argues that it is desirable to analyze dependent indefinites and adjectives with a sentence-internal reading, like *same* and *different*, as a unified class of phenomena. His concern for a scalar implicature approach is that it lacks generality: while it may be a reasonable account for dependent indefinites and binominal *each*, it cannot be extended to adjectives with a sentence-internal reading. I do not pursue an implicature analysis in terms of competing scope configurations as it fails to account for the counting quantifier requirement and the monotonic measurement requirement.

[9]The term counting quantifier does not have an agreed-upon definition in linguistics. For example, while *few* is taken to be a counting quantifier in Beghelli and Stowell (1997), it is not treated as one in Sutton (1993).

[10]To some speakers, *some* and *few* are better than the rest in (12) (Simon Charlow, p.c.). Safir and Stowell (1988) suggest that *some of the NPs* are slightly more acceptable than *some NPs* when they serve as the host. They judge the latter as unacceptable. I have omitted the singular indefinite determiner *a* from (12) as it exhibits inter-speaker variability. I will return to the inter-speaker variability of indefinites involving *a* in section 3.3.

(11) The girls saw $\left\{\begin{array}{c}\text{two}\\\text{at least two}\\\text{more than two}\\\text{a few}\\\text{several}\\\text{many}\\\text{a lot of}\end{array}\right\}$ movies each.

(12) *The girls saw $\left\{\begin{array}{c}\emptyset\\\text{some}\\\text{a certain}\\\text{the}\\\text{those}\\\text{few}\\\text{most}\\\text{all}\end{array}\right\}$ movie(s) each.

Most previous studies that take up the counting quantifier requirement take it to reflect a need for indefinites or number terms. For example, Zimmermann (2002) takes binominal *each* to be only composable with an indefinite. However, this analysis over-generates, as many indefinites in (12) cannot host binominal *each*, such as *some movies* and *a certain movie*. In fact, many speakers even dislike regular indefinites with the determiner *a* for hosting binominal *each*, according to Safir and Stowell (1988:(7a)):

(13) %The men saw a jewel each.

Cable (2014) tries to reduce the counting quantifier requirement to a subcateogorization for number terms. This has the effect of ruling out the hostile hosts in (12) as they do not contain any number term. There are two problems with this type of reduction. At an empirical level, it does not straightforwardly explain why the indefinite article becomes more acceptable when it is part of a counting quantifier or a measure phrase, as in (15).

(14) The men saw a few jewels each.

(15) The children ate a bag of candy each.

Postponing to section 3.3 a more in-depth analysis of the contrasting roles of the indefinite article in (13) on the one hand and (14) and (15) on the other hand, what is worth highlighting here is that the counting quantifier requirement does not follow merely from a subcategorization rule against the indefinite article *a*.[11]

At a theoretical level, even if it is empirically adequate to explain the counting quantifier requirement as a need for number(-like) terms, it still raises the question *why* number terms matter at all for a distributivity marker. Surprisingly, an answer to this question does not come from number terms themselves, but from a component intimately related to them, namely, *measure functions*. In the next subsection, I show that binominal *each* is sensitive to the type of measure function associated with its host. Once the measurement sensitivity is in place,

---

[11]The challenge being mounted here is not fool-proof. It can be explained away by assuming that what binominal *each* subcategorizes for is a degree quantifier, rather than a degree or a number. It is possible to treat number terms like *three* and *five*, counting phrases like *a few*, and measure phrases like *a bag of* as degree quantifiers along the lines of Kennedy (2015).

the counting quantifier requirement follows as all counting quantifiers are associated with an appropriate measure function (i.e., cardinality).

2.3   From monotonic measurement to monotonicity in relation to distributivity

Zhang (2013) is the first to observe that the type of measurement plays a crucial role in constraining what noun phrases may host binominal *each*: monotonic measurements give rise to friendly hosts but non-monotonic measurements give rise to hostile hosts. Postponing a review of Zhang's own account until section 5, let me give weight to Zhang's observation by offering some background on the semantics of measurement.

In the literature on measurement, it is widely assumed that numeral expressions such as *two students*, *seven feet*, and *30 miles per hour* have more structure than meets the eye. In addition to the numeral and the common noun, they also contain measure functions such as cardinality, height, and speed. Many studies assume that measure functions are either directly syntactically represented in certain noun phrases (e.g., Krifka 1995) or they are contextually associated with functional categories that are syntactically represented in certain noun phrases (e.g., Hackl 2000, Schwarzschild 2006). For the purpose of this study, I assume that measure functions are syntactically represented in noun phrases like counting quantifiers and measure phrases.

According to Lønning (1987), a measure function denotes a mapping from a class of physical objects to a degree that preserves a certain empirically given ordering relation, such as "be heavier than" (for weight) or "be cooler than" (for temperature). Degrees are further mapped to numbers by unit functions like pound or kilogram.

Schwarzschild (2006) and Wellwood (2015) further classify measure functions into two types—monotonic and non-monotonic measure functions.[12] These two types of measure functions differ with respect to the relationship between an object domain and a corresponding a degree domain. More concretely, a monotonic measure function like weight preserves the part-whole structure in the two domains, so that the weight of the entire object is always bigger than the weight of any proper part of the object; whereas temperature is non-monotonic since the temperature of a divisible object is *not* always bigger than the temperature of any of its proper parts.

The examples in (16) and (17) demonstrate Zhang's observation that binominal *each* can only be hosted by a noun phrase with a monotonic measure function. All the measure functions in (16) are monotonic while the ones in (17) are non-monotonic.[13]

(16)   Monotonic measurement

      a.   The boys read two books each.   cardinality

---

[12]A closely related characterization of monotonic measure functions is extensivity (Krifka 1989). However, while extensivity requires additivity, as defined in (i) (a simplified version based on Krifka 1989: D24), monotonicity does not, as shown in (24), and hence is weaker than the extensivity.

(i)   A measure function μ is additivity iff $\forall x \forall y : \mu(x) + \mu(y) = \mu(x \oplus y)$

As far as I know, there is no convincing empirical argument on whether the right notion for characterizing measure functions in natural language should be extensivity or monotonicity. Although both terminology exist to characterize the same set of measure functions, I opt for the term *monotonicity* in this paper as this weaker notion is a better characterization of the properties of binominal *each*.

[13]While all the sentences in (16) are judged as acceptable, not all sentences in (17) are judged as unacceptable by the five native speakers I consulted. In particular, (18-a) received an average rating of 2.2, (87) an average rating of 2.4, and (17-c) an average rating of 2.8. All of the ratings are based on a 5-point Likert scale, with 1 labeled as 'completely unacceptable', 3 labeled as 'marginal', and 5 labeled as 'fully acceptable'.

   b. The girls walked three miles each.   distance
   c. The windows are four feet (tall) each.   height

(17) Non-monotonic measurement

   a. *The drinks are 60 degrees (Fahrenheit) each. temperature
   b. *The girls walked at three miles-per-hour each.   speed
   c. *The gold rings are 24 Karat each.   purity

Let me address three potential concerns here. The first one is a concern about form. Some of the non-monotonic measure phrases seem more complex than the monotonic ones, such as the case with (87), or less frequent, as in the case of (17-c). To sharpen the monotonic measurement requirement, it is desirable to construct a minimal pair. The two examples in (18) both have a host in the form of a measure phrase *60 degrees* and serve this purpose. The one associated with temperature (which is non-monotonic) is still unacceptable while the one associated with angle degrees (which is monotonic) becomes acceptable.[14]

(18) a. *The drinks are 60 degrees (Fahrenheit) each. temperature
   b. The angles are 60 degrees each.   angle

Another concern is that measurements like cardinality, distance, and height may be in some sense more basic (or concrete) than measurements like temperature, speed, and purity. So, perhaps it is the basic-ness of measurement that binominal *each* cares about rather than the monotonicity distinction. Comparing examples like (19-a) on the one hand and examples like (19-b) and (19-c) (provided by a reviewer) can help address this concern.

(19) a. *The drinks are 60 degrees (Fahrenheit) each. temperature
   b. The naughty children raised the living room temperature by two degrees each.
                   temperature change
   c. The drops of hot water raised the temperature of the solution by two degrees each.
                   temperature change

The measure function associated with the host in (19-a) is temperature. Since it does not track the massiveness of the object being measured, it is a non-monotonic measure function. By contrast, the measure function associated with the host in (19-b) and (19-c) is temperature change. Unlike temperature, temperature change is additive—if a child raised the temperature two degrees and another child raised it two more degrees, then together they raised the temperature four degrees. Since there is no reason to believe that the measure function temperature change is in any way more basic or concrete than the measure function temperature, I take the contrast in the examples in (19) to lie in the monotonicity distinction of the measure functions associated with the hosts.

  The last concern is partly related to the second one. One may think that expressions like *two books* and *three miles* can be associated with distinct entities (as they could be argued to be more concrete in some sense). After all, one can easily talk about distinct sets of two books and distinct 3-mile paths. By contrast, expressions like *60 degrees* and *24 Karat* cannot be easily associated with distinct entities. What does it mean to have two distinct 60-degree temperatures

---

[14]For (18-b) to be fully acceptable, some speakers prefer intersections placed close to each other to form adjacent angles or arranged to form internal angles of a polygon, while others, such as one of the reviewers, are fine with a random arrangement. It is not clear to me what underlies the inter-speaker variation. To the extent that arrangement matters, it provides some support for the claim that measurement of a collection is involved, as previous studies have shown that arrangement is a factor determining measurement of collections (Scontras and Goodman 2017).

and two distinct 24-Karat purities? However, the fact that (19-b) and (19-c) are well-formed shows that individual variation does not really matter. *Two degrees* as a measurement of temperature change is also hard to associate with distinct entities—to the extent that it makes sense to talk about two distinct 2-degree changes in temperature, one is inherently thinking in terms of additivity, which is a defining property of monotonic measurement.

Zhang's observation about measurement sensitivity is an important one as it allows us to make sense of the counting quantifier requirement without stipulating it as a syntactic, subcategorization requirement—counting quantifiers like *two movies*, *a few books*, and *many chairs* are special because they are all associated with a particular monotonic measure function, namely, cardinality. As a result, the counting quantifier constraint is reducible to the need for a monotonic measure function.

Zooming out of distributivity, binominal *each* is not the only natural language expression to which the distinction between monotonic and non-monotonic measurement matters. Schwarzschild (2002, 2006) points out a similar contrast in pseudo-partitives: pseudo-partitives admit monotonic measurement, as in (20), but reject non-monotonic measurement, as in (21).

| (20) | a. | two pounds of cherries | weight |
| | b. | thirty liters of water | volume |

| (21) | a. | *five degrees Celsius of the water in this bottle | temperature |
| | b. | *five miles an hour of running | speed |

Interestingly, temperature change can occur in pseudopartitives, which only admit monotonic measure functions. Consider the following example, due to Champollion (2017).

(22) The scientists from Princeton and Harvard universities say just two degrees Celsius of global warming, which is widely expected to occur in coming decades, could be enough to inundate the planet. (Champollion 2017:141)

Given the monotonic nature of the measure function temperature change, it is hence not surprising that it can support the use of binominal *each*.

In addition, Wellwood (2015) observes similar contrasts in comparatives. Both sentences in (23) can express comparisons involving monotonic measurement, but neither can express a comparison involving non-monotonic measurement. For example, in (23-a) the amount of the soup that Al bought is larger than the amount of the soup that Bill bought. The amount may be understood in terms of volume or weight, but not temperature.

| (23) | a. | Al bought as much soup as Bill did. | volume, weight, *temperature |
| | b. | Al ran as much as Bill did. | time, distance, *speed |

Schwarzschild (2002, 2006) accounts for the sensitivity of measurement constructions to types of measure function by invoking the notion of **monotonicity**. Wellwood (2015) provides a formal definition of this monotonicity condition on measurement, as shown in (24). This condition requires that the part-whole structure of the domain of a measure function be strictly preserved in the domain of degrees.

(24) **Monotonic Measurement** (Wellwood 2015)
A measure function μ is monotonic iff

   a.   there exists $x, y \in D_{\sqsubseteq^{part}}$, such that $x \neq y$, and
   b.   for all $x, y \in D_{\sqsubseteq^{part}}$, if $x \sqsubset^{part} y$, then $\mu\ x <^{deg} \mu\ y$

The monotonicity condition of Schwarzschild and Wellwood requires that only monotonic measure functions can be used in measurement constructions like pseudo-partitives or comparatives. Consequently, monotonic measure functions, but not non-monotonic ones, pass the condition. I argue that at the core of binominal *each* is also a form of measurement-related monotonicity condition, but one that makes critical use of distributivity, unlike the monotonicity condition on measurement constructions, which does not appeal to distributivity. The rest of this paper is devoted to fleshing out such a monotonicity condition, explaining its relationship to distributive quantification, and extending it to account for the three requirements discussed in this section.

3 A monotonicity constraint for binominal *each*, informally

I propose that a sentence with binominal *each* makes a two-part contribution: distributivity and monotonicity. The former may be contributed by binominal *each* itself, as argued in many other studies that take *each* to be a distributive operator (e.g., Zimmermann 2002, Dotlačil 2012, Champollion 2017), or it may be contributed by a separate distributive operator, as suggested in Champollion (2015) following Henderson's (2014) semantics for dependent indefinites. I leave both options open as there has been relatively little work on the (potentially degraded) acceptability of binominal *each* when a distributive quantifier is present, as exemplified in (25) (see also related discussions in Champollion 2015 and Kuhn 2017).[15] Generally speaking, the less compatible binominal *each* is with a distributive quantifier, the more appealing is the claim that it contributes a distributive operator. For the sake of concreteness, the semantics developed for binominal *each* in this paper assumes the second option, namely, that the distributive inference is contributed by a separate distributive operator rather than by binominal *each*.[16] It is perhaps worth emphasizing that the first option can also be pursued with some minor changes to the lexical entry of binominal *each* (a lexical entry for the 'distributive-version' of binominal *each* is given in footnote 39 for concreteness).

(25)   ?Every girl saw two movies each.

What is important for us is that the distributivity component establish a dependency between values introduced by the Key and values introduced by the host. For example, (26) allows us to relate the girls to the movies they saw, as indicated by the left dependency in Figure 4. The dependency is notated as f, which can be thought of as a function that maps each girl to the movies she saw and also sums of girls to the sums of movies they saw. In other words, f encodes the functional dependency induced by distributivity and is cumulatively closed (marked by *, following Link 1983). This dependency, whose formal formulation will be given in section 4, provides an important ingredient for formulating the monotonicity condition for binominal *each*.

(26)   The girls saw two movies each.

---

[15]It was brought to my attention that Bauman et al. (2012) reported middling judgements for sentences like (25). However, I was unable to secure the work for evaluation. The sentence is marked with '?' instead of '%' to indicate its potential degradedness, following the suggestion from a reviewer.

[16]There are two nontrivial consequences from not attributing distributivity to binominal *each*. First, it remains unexplained why binominal *each* is not readily compatible with distributive quantifiers, as pointed out by a reviewer of this paper. Second, it needs to be explained why the distributive interpretation, but not the collective or cumulative interpretation, licenses the use of binominal *each*.
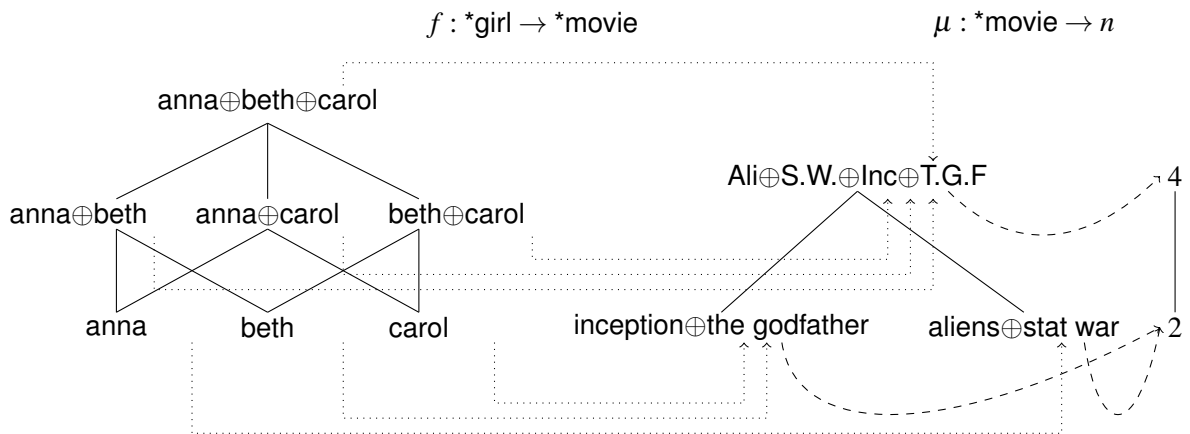
Figure 4: A dependency arising from distributive quantification.

Another piece of ingredient needed to assemble the monotonicity constraint is a measure function μ. What μ does is map the individuals contributed by the host to a degree (which can be modeled as a point on a scale). For example, the cardinality measure function associated with *two movies* in (26) maps the movies each girl saw to the number 2 and the movies they collectively saw to the number 4, as indicated by the mapping on the right in Figure 4.

Together f and μ form a composite mapping from the the plurality being distributively quantified to the degrees associated with the size of the host.[17]

A monotonicity constraint in association with distributivity is precisely a requirement that the composite mapping be monotonic. That is, the range of μ should preserve the part-whole relation found in the domain of f. For concreteness, a framework-neutral monotonicity constraint is formulated in (27).

(27) **Monotonicity in association with (the mereological structure of) distributivity**
A measure function μ is a monotonic measurement in association with f iff the following two mappings hold:

a. Non-constant mapping
There are distinct $a, a' \in \mathsf{Dom}\ f.\ \mu(f\ a) \neq \mu(f\ a')$

b. Non-decreasing mapping
For all $a, a' \in \mathsf{Dom}(f).\ a \leq a' \rightarrow \mu(fa) \leqslant \mu(f\ a')$

---

[17]In fact, the mapping can be thought of as the composite function $\mu \circ f$, constructed with help of the two mappings f and μ. A reviewer asked if the composition function $\mu \circ f$ may itself be seen as a measure function. That is, instead of simple measure functions like weight or height, which measure people's weight or height, $\mu \circ f$ measures people in terms of movies watched. As also pointed out by the reviewer, such a move has the benefit of rendering (27) as essentially identical to (24). I think this is a correct paraphrase of what the composition function is formally doing—it takes any plural individual and maps all its parts to the measurement of its host. However, I do not think that it is conceptually appealing. For one thing, there is no reason why measurement of a plurality has to involve distributivity. A simple plurality with a part-whole structure will suffice, as witnessed in sentences like *The books weigh 10 pounds*—there is no need to bundle distributivity and measurement. However, there is a way of thinking $\mu \circ f$ in terms of a complex measure function that is conceptually quite appealing and is intimately related to the reviewer's suggestion. It is to think of it as measuring some dimension of a distributive dependency. A dependency is essentially a multi-dimension object, in the same way that an event is. This object may be measured on its Key dimension (i.e., the dimension representing all the values introduced by the Key) or on its Share dimension (i.e., a dimension that depends on the Key). Such a move requires us to think of dependencies as concrete entities that can be measured, in much the same way that individuals paired with events or situations can be counted like concrete individuals in studies such as Krifka (1991) and Barker (1999).

Let me elaborate on the two mappings that make up the monotonicity constraint. The non-constant mapping in (27-a) essentially requires *measurement variability* in the range of f. It partially resembles the domain plurality clause found in the definition of monotonic measure functions in (24-a). Both require that there are distinct individuals in the domain of a function. In the domain plurality clause in (24-a), the requirement is imposed on the domain of the measure function. However, in the non-constant mapping clause in (27), the requirement is not imposed on the domain of the measure function, but on the domain of f, which is the plurality being distributively quantified. The non-constant mapping clause also has an additional requirement of measurement variability in range of f. What this amounts to is that the host must exhibit variability in measurement in association with the plurality being distributively quantified. As will be shown shortly, non-constant mapping is useful for capturing the variation requirement of binominal *each*.

The non-decreasing mapping in (27-b) requires that the part-whole relation found in the domain of f (i.e., the girls) be mapped *non-decreasingly* to the measurement of the range of f (i.e., the movies). Since f encodes a dependency induced by distributivity, this amounts to making reference to the *mereological structure* of distributivity. In other words, we are referring to parts of a dependency that stand in a part-whole relation. Modulo the association with f, (27-b) is a standard definition of non-decreasing monotone functions. It is weaker than the definition of monotonic measure functions found in Wellwood (2015) as cited in (24-b), which picks out strictly increasing functions among the non-decreasing ones.[18]

3.1 Capturing the variation requirement

Recall from the discussion in section 2.1 that the variation requirement requires that the host of binominal *each* co-varies with the Key. For example, (28) requires the girls to not all see the same two movies.

(28)   The girls saw two movies each.

The monotonic measurement condition stated in (27) straightforwardly captures the variation requirement. The crucial piece comes from (27-a), which requires that there are at least two distinct values in the domain of f and they are associated with distinct measurements of the host. If the distributivity contributes a dependency like the one in Figure 4, (27-a) is satisfied: there are two distinct individuals, for example, anna and anna⊕beth, such that they saw different *numbers* of movies. Concretely, the atomic individuals saw two and the plural individual

---

[18]The decision on a weaker form of monotonicity, one in terms of a non-decreasing mapping, instead of a stronger form requiring a strictly increasing mapping, as suggested in Wellwood (2015) for monotonic measure functions and adapted below in (i) to incorporate association with distributivity, is empirically motivated. If a strictly increasing mapping is required, then in (26) no pairs of girls can watch the same two movies, contrary to fact, which only requires at least two girls to see different movies. In other words, the fact that the composite function in Figure 4 is acceptable suggests that the mapping should be non-decreasing rather than strictly increasing.

(i)   **Strictly increasing dm** (rejected)
        For all $a, a' \in$ Dom f. $a < a' \rightarrow \mu(f\,a) < \mu(f\,a')$

On a related note, a reviewer asks if a case could be made that perhaps all measure functions characterized as involving a strictly increasing function in Wellwood (2015) actually involve a weaker, monotonic (non-decreasing and non-constant) function. While I hold the same suspicion, it is not easy to test the hypothesis. To get the hypothesis tested, one would need to find measure functions that are monotonic but not strictly increasing and test if they pattern differently from their strictly increasing peers in linguistically meaningful ways. So far, I have not found satisfactory examples for such measure functions.

saw four.[19] However, if distributivity contributes a dependency that maps all the girls to the same two movies, then the condition cannot be satisfied: there are no girls (atomic or plural) that saw different numbers of movies.

At first glance, capturing the variation requirement, which is originally stated as variation in individual identity, in terms of measurement variability may seem like merely a clever trick. Adding up different individuals of course creates a more massive plural individual and hence yields a larger measurement. However, measurement variability is more adequate than individual variation as an explanation for the variation requirement when hosts beyond counting quantifiers are considered.

In particular, measure phrases can also serve as a host for binominal *each*, as shown in (29).

(29)　　a.　The angles are 60 degrees each.
　　　　b.　The naughty children raised the living room temperature by two degrees each.

In (29-a), the host does not contribute any individuals other than the angle degrees. So, it is hard to talk about individual variation. Moreover, there is no variability in the degree associated with each angle as all of the angles are asserted to be 60 degrees. The only way to achieve variation is by considering the mereological structure of the angles and how it stands in relation to the angle measurements. Recall that measurement variability does not require the host to be distinct individuals in association with distributivity. It merely requires the host to yield distinct *measurements* in association with the *mereological structure* of distributivity. Since concatenating two smaller angles (along a certain dimension) does yield a straightly larger angle, measurement variability can be achieved as long as there are at least two distinct angles being considered.

In (29-b), the measure phrase also does not contribute any individuals.[20] Again, the temperature that each child adjusted is the same, i.e., two degrees. The variation requirement is only satisfied because the mapping from the children plurality to the degrees of temperature change is non-constant.

3.2　Capturing the monotonic measurement requirement

Let me turn next to showing how the monotonic measurement requirement is captured by the monotonicity condition in (27). To begin with, it is useful to return to a sentence with a counting quantifier host such as (26). The composite function $\mu \circ f$ is illustrated in Figure 4. The measure function in this case is cardinality (or μ_{card}). As discussed in the previous subsection, $\mu \circ f$ satisfies measurement variability as stated in (27-a). In addition, $\mu \circ f$ also satisfies the non-decreasing mapping as stated in (27-b). To see this, we need to check that none of the bigger elements in the domain of $\mu \circ f$ is mapped to a smaller degree in the range of the function. Since cardinality is a monotonic measure function, a bigger number of girls will never saw a smaller number of movies, our mapping is guaranteed to be non-decreasing.

Having seen how the monotonic measurement condition correctly rule in counting quantifier hosts, we turn next to how the condition distinguishes between monotonic and non-monotonic measure phrases. We first consider measure phrases with an unmodified numeral, such as the ones in (30-a) and (30-b).

---

[19]If we compare the number of movies watched by each girl, there is no variation in the measures. The variation in measures only arises when the girls form a part-whole structure and hence their dependent measures also form a part-whole structure.

[20]Even if we take the measure phrase to contribute the air particles in the living room, there can be no variation in these particles when each child raised the thermostat and the sentence is still acceptable.

(30) a. *The coffees are 60 degrees each.   temperature
      b. The angles are 60 degrees each.   angle

Since the measure phrases directly measure the entities denoted by subjects, let us assume that the dependency f is trivial and maps any entities to themselves in the examples above. This assumption is later reflected in the definition of binominal *each* hosted by measure phrases in (60-b). μ still serves to relate coffees (or angles) to their corresponding temperature (or angle degrees). Composing these two sets of functions yields the two composite functions in as visualized in Figure 5.

$\mu \circ f : {}^*\text{coffee} \to D_d$ $\mu \circ f : {}^*\text{angle} \to D_d$

$c_1 \oplus c_2$ $a_1 \oplus a_2$ 120

                          60

c_1 c_2 a_1 a_2 60

(a) Temperature   (b) Angle degree
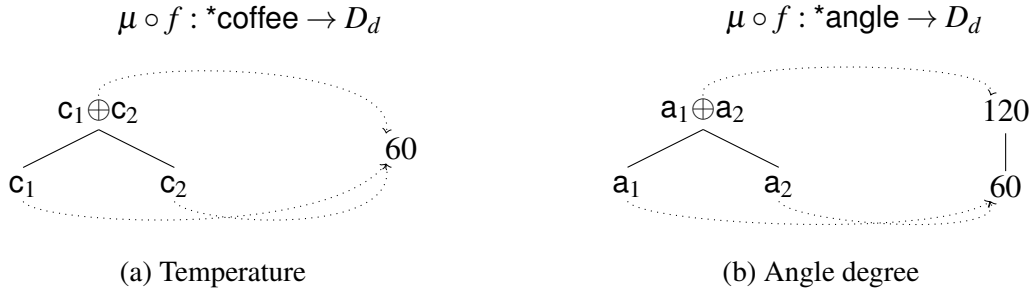
Figure 5: Non-monotonic vs. monotonic measurement

For concreteness, the coffees (or angles) are distributively checked for their temperature (or angle degree). f encodes a functional dependency between coffees (or angles) and themselves. μ_{temp} (or μ_{ang}) is a temperature (or angle) measure function. The subsequent function resulting from composing f and μ is hence identical to μ. Both sentences satisfy non-decreasing mapping as stated in (27-b). However, (30-a) fails non-constant mapping as stated in (27-a) while (30-b) satisfies it. This is because measurement variability is missing in (30-a), as demonstrated in 5a, but present in (30-b), as shown in Figure 5b.[21]

One may suspect that (27-a) alone is sufficient to guarantee the variation requirement and the privilege of monotonic measure functions. It is not. It can be satisfied with an non-monotonic measure function as long as the function yields different degrees for different values in the range of f. For example, consider binominal *each* whose host is a measure phrase with a *modified numeral*, such as (31-a) and (31-b). Figure 6 illustrates how f and μ_{dim} work in these two sentences.

(31) a. *The drinks are more than 60 degrees each.

---

[21]While I am assuming here that it is defined to measure the temperature of a plural entity, this should not be taken for granted. In many situations, the 'collective' temperature (or speed, or other measures from non-monotonic measures) may not even be defined. If this is the case, then the monotonicity condition may be violated simply because it is impossible to measure a pluralities on certain dimensions. However, it is not entirely clear to me that non-monotonic measurement of pluralities is always undefined. We can talk about the velocity of a river despite the fact that different parts of the river flow at different speeds. Similarly, we can talk about the temperature of the earth, despite the fact that different parts of the earth have distinct temperatures.

Relatedly, it has been brought to my attention that measurement of pluralities may be represented as degree pluralities along the lines of Dotlačil and Nouwen (2016) (see also Schwarzschild 2006, Beck 2010), which are essentially collections of degrees (much like collections of individuals). As far as I can see, this approach to measurement of pluralities is compatible with the present analysis, as long as binominal *each* is taken to enforce some form of summation of two degrees. While measures from monotonic measure functions sum up easily (as they track the bulk of the entities being measured, which themselves can be summed), measures from non-monotonic measure functions either do not sum up (giving rise to undefinedness) or may sum up in some way but violates monotonicity. I thank a reviewer for drawing my attention to degree pluralities and its relevance to the present study.
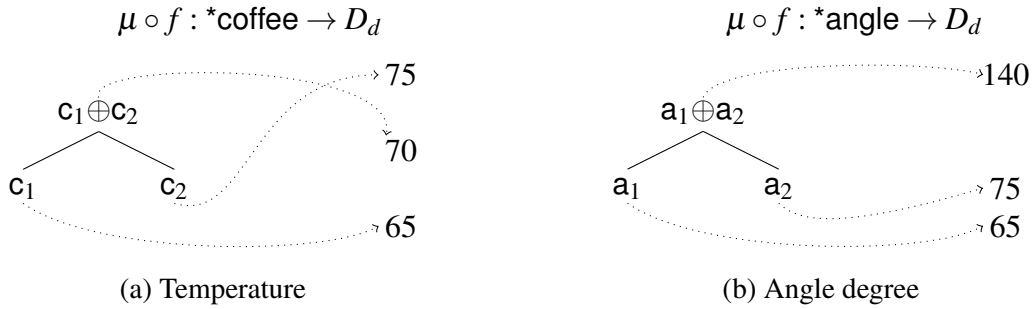
b.   The angles are more than 60 degrees each.

$$\mu \circ f : {}^{*}\text{coffee} \rightarrow D_d \qquad\qquad \mu \circ f : {}^{*}\text{angle} \rightarrow D_d$$

(a) Temperature　　　　　　　　　　　(b) Angle degree

Figure 6: Modified numerals interacting with monotonic vs. non-monotonic measurement

(31-a) satisfies non-constant mapping, as the range of f has different degrees. However, it is still not well-formed. This is because it violates non-decreasing mapping: there is a pair of elements in the domain of f that stand in a part-whole relation whose corresponding measurement fails to preserve the the order of the pair, as indicated by the crossing lines in Figure 6a. By contrast, (31-b) satisfies both non-constant mapping and non-decreasing mapping, as indicated in Figure 6b.

3.3   Capturing the counting quantifier requirement

Lastly, we predict that noun phrases without an appropriate measure function component cannot host binominal *each*. A natural question that arises is how we can diagnose the presence of a measure function component. I do not have a comprehensive answer to this question. However, compatibility with unit words like *pound(s)* and *mile(s)* seems to be a rather reliable test: if a determiner-like expression is compatible with these expressions, then it can form a noun phrase that can host binominal *each*.[22] Some examples are given in Table 1.[23] Included with each type of expression are studies that attribute measure functions to it. There is no study that attributes measure functions to quantificational determiners except for *few* (Schwarzschild 2006) and *most* (Solt 2016). However, both *few* and *most* fail to pass the unit word test, as demonstrated in (33). For this reason, I take them to be incompatible with the measurement component that may go with (modified) numerals and quantity expressions.[24]

(32)   *The girls biked most miles each.

---

[22]The numeral term *zero* is predicted to be incompatible with binominal *each* even if it occurs in an appropriate counting quantifier or measure phrase such as *zero books* and *zero miles*. For some (but not all) speakers, this prediction is borne out. It's unclear me to what underlies the individual variation.

[23]*Some* sometimes does occur with unit functions, as in *gained some inches* and *lost some pounds*. In these cases, the unit functions are interpreted as standing in for the entities they measure, i.e., *height* and *weight*, respectively. I have been informed that *some + units* are more friendly hosts than ordinary *some NPs* (Simon Charlow, p.c.):

(i)   The boys lost some pounds each over the summer.

(ii)   *The boys lost some marbles each over the summer.

[24]A reviewer observes that *most miles* in (33-b) is improved in examples like *Paul biked most miles on Monday. (Only 20 were left to go on Tuesday.)* So, its unacceptability in (33-b) may be due to other reasons. I suspect what happens in the improved example is that *miles* gets to stand in for something like *the distance (in miles)*. Note that even in the improved environment, *most miles* still cannot host binominal *each*:

Expressions	compatibility with unit words	host binominal *each*
(modified) numerals two, at least/most two, more/less than two e.g., Hackl (2000), Kennedy (2015)	yes e.g., two pounds more than five miles	yes e.g., two books each
quantity expressions a few, a couple, many e.g., Rett (2014), Solt (2015)	yes e.g., a few gallons	yes e.g., many movies each
quantity comparatives more, as many (much) as e.g., Wellwood (2015)	yes e.g., as many pounds as	yes as many books each
quantificational determiners no, some, few, most, every, all	no e.g., *most miles	no e.g., *most books each

Table 1: Expressions that can (and cannot) host binominal *each*

(33) a. Peter gained many/??few/*most pounds over the summer.
 b. Paul biked many/??few/*most miles every day.

It has been noted that noun phrases with the indefinite article *a* are better than those with the determiner *some* in hosting binominal *each*, although not all speakers accept the former (Safir and Stowell 1988, Szabolcsi 2010, Milačić et al. 2015), as illustrated in (34).

(34) a. %The boys read a book each.
 b. *The boys read some book(s) each.

Interestingly, *a* is also compatible with measurement units in ways that *some* is not, as shown in (35). This is not at all surprising given that many linguists have argued that *a* is derived from *one* synchronically and/or diachronically (e.g., Perlmutter 1970, Chierchia 2013, Kayne 2015). Given these considerations, it is conceivable that *a* is ambiguous between a numeral *one* and an existential determiner, while *some* is only an existential quantificational determiner without a measure function component.

(35) a. Julia ran a/*some mile yesterday.
 b. Peter pumped a gallon/*some gallon (of gasoline) into his car last week.
 c. The sea level has risen an/*some inch over the last three months.

Note that *some* is not entirely incompatible with a measurement unit. The compatibility improves when the measurement unit is in the plural form, as shown below:

(36) a. Julia ran some miles yesterday.
 b. Peter pumped some gallons of gasoline into his car last week.
 c. The sea level has risen some inches over the last three months.

I will not go into why number morphology makes a difference to the compatibility of *some* and measurement units.[25] However, the fact the *some* is compatible with measurement units can be used to further corroborate the relevance of measurement in licensing binominal *each*. When

---

[25]It has been observed that when a singular noun follows *some*, the noun phrase exhibits some kind of epistemic effect. It is possible that that epistemic effect is incompatible with measurement units.

a measurement unit occurs in a *some NP*, the *some NP* can host binominal *each*, as shown in (37).

(37)   a.   The girls ran some miles each yesterday.
            b.   The drivers pumped some gallons of gasoline each before the electricity went out.

4 Formalizing the monotonicity condition

Now that the monotonicity condition in association with distributivity has been established, we are ready to embed it in a compositional semantics. It is not difficult to see what kind of framework we need to implement this condition compositionally. An adequate framework should satisfy the following three criteria:

- Criterion 1: It should allow us to talk about plurality and measure functions of various sorts.

- Criterion 2: It should be able to represent the functional dependency arising from distributive quantification. In other words, it should make concrete how f is constructed.

- Criterion 3: Since the monotonicity constraint checks the measurement of the host in association with a distributive dependency, the dependency has to be established before the monotonicity constraint can be checked. For this reason, it is necessary to split up the contribution of a host of binominal *each*, evaluating one part (i.e., it basic semantics as a counting quantifier or a measure phrase) inside the scope of distributivity and the other part (i.e., the monotonic measurement condition) outside the scope of distributivity. The former provides the necessary ingredients for building the functional dependencies of distributivity and hence the function f. The latter can access f after it is assembled.

Criterion 1 is very easy to satisfy. Any framework that can be enriched to include pluralities and measure functions can be used to model monotonicity. Therefore, a decisive choice depends on the remaining two criteria.

There are a few frameworks that satisfy Criterion 2, including Krifka's (1996) framework of DPL extended with parameterized sum individuals, van den Berg's framework of Dynamic Plural (Predicate) Logic (DPlL), which is also an extension of DPL, Brasoveanu's (2008) framework of Plural Compositional DRT. This is not an exhaustive list. Event-based distributivity and Skolem function-based distributivity (or universal quantification) can also be used to generate dependencies as a result as evaluating distributivity. See Schein (1993), Lasersohn (1995) and Champollion (2017) for the former and Huang (1996) for the latter.[26] In this paper, I sketch a variant of DPlL that borrows features from both the original DPlL and PCDRT. The merit of a DPlL/PCDRT-type logic is that it not only tracks the dependency in context, but the built-in anaphoric device (i.e., discourse variables) allows us to access the dependency relatively easily. This type of logic has been used to model phenomena that require reference to the functional dependency of distributivity, such as quantificational subordination (van den Berg 1996, Nouwen 2003), quantifier-internal anaphora and reciprocals (Dotlačil 2010), as well as dependent indefinites (Henderson 2014, Champollion 2015 and Kuhn 2017). It is not surprising

---

[26]It is also worth mentioning that the linking function f that is crucial in the formulation of the monotonicity condition (e.g., (27)) bears striking resemblance to thematic functions in event semantics—a remark credited to one of the reviewers of this paper.

that the same type of logic can be extended to capture the monotonicity constraint of binominal *each*, as reference to the mereological structure of a dependency is merely a subtype of reference to dependencies.

Criterion 3 essentially asks for a split-scope mechanism. Several alternatives have been explored in the literature of dependent indefinites, including the use of post-suppositions (Henderson 2014), standard scope (Kuhn 2017), and scope of higher-order meanings (Charlow, to appear). This paper opts for the use of higher-order meaning, following the empirical motivations discussed in Charlow (to appear).

It should be clear by now what kind of framework is needed to account for the novel properties of dependent indefinites observed in this work. In the next sections, the essential components of such a framework are provided. I begin by discussing the general framework in section 4.1, followed by translating the monotonic measurement condition into this framework, and lastly in section 4.3 the monotonicity condition is implemented in a compositional manner.

4.1 Formal background: DPlL with Referential Pluralities

The direct logical predecessors of the framework used in this paper are van den Berg's (1996) DPlL and Brasoveanu's (2008) PCDRT. Like DPlL, it assumes that dependencies do not arise without distributivity; like PCDRT, it assumes referential pluralities (i.e., that the domain of individuals have both singular and plural individuals) and sub-sentential compositionality. These design features are elaborated upon when they get introduced later in this section. To distinguish the present logical framework from its predecessors, it is referred to as 'DPlL with Referential Pluralities'. Only the core logical backbone and essential apparatuses are introduced in this section. For a more comprehensive version of this logic, see Chapter 2 of Law (2019).[27] Readers who are familiar with the difference between DPlL and PCDRT can treat this section as a formal appendix with the exception of Definitions 9 and 10.

4.1.1 Dynamics and plural info-states

Like both predecessors, the dynamic core of the present logical framework is built upon Dynamic Predicate Logic (DPL, Groenendijk and Stokhof 1991). As such, it follows the spirit of DPL and takes the interpretation of a sentence to be a *relation between contexts*. In DPL, a context is a *single assignment*, so a sentence is said to be evaluated to true ($\mathbb{T}$) relative to an input assignment g and and output assignment h (and a model M, omitted for readability) if and only if they belong in the relation denoted by the sentence:

(38) $g[\![\phi]\!]h = \mathbb{T}$ iff $\langle g, h \rangle \in [\![\phi]\!]$[28]

Superimposing a plural logic on DPL yields DPlL. This is done by blowing up the size of a context from a single assignment to a *set* of assignments, known as **information states** (or info-states for short), as defined below:

Definition 1 (Information state)
An information state (info-state) is a set of assignments.

For readers who are interested in DPlL/PCDRT-type semantics, it is recommended that they start with DPL (Groenendijk and Stokhof 1991), since DPlL/PCDRT is built upon DPL. Detailed introductions to DPlL/PCDRT are found in dissertation studies (because they do not usually have length limitations), such as van den Berg (1996), Brasoveanu (2007), Dotlačil (2010), Henderson (2014), Kuhn (2015), DeVries (2016), and Law (2019).

[28]Building on this, a sentence ϕ is true relative to an input assignment g iff there is an output assignment h such that $g[\![\phi]\!]h = \mathbb{T}$. This notion of truth is useful for comparing truth in a dynamic framework and a static framework.

A sentence in DPlL then denotes a relation between info-states. In other words, fixing an input info-state and an output info-state determines the truth value of a sentence:

(39) $G[\![\phi]\!]H = \mathbb{T}$ iff $\langle G,H \rangle \in [\![\phi]\!]$

Following a common practice in the literature, an information state is represented as a matrix, as shown in Figure 7. The first row houses the variables introduced into the info-state so far while the first column lists all the assignments in the info-sate . Starting from the second column, each column represents the collective value associated with a specific variable. Starting from the second row, each row represents an *n*-tuple of values associated with different variables. The entire matrix represents **dependencies** among variables, which are precisely what we need for constructing a monotonicity condition in association with a distributive dependency.

G	...	x	y	...
g_1	...	a	d	...
g_2	...	b	e$\oplus$f	...
g_3	...	c	g	...
...	...	...	...	...

Figure 7: A sample info-state

In the original DPlL logic, although a collection of assignments may map a variable to a set of values, a single assignment can only map a variable to a single value. This restriction is relaxed in its cousin logic, Plural Compositional DRT (PCDRT, Brasoveanu 2007, 2008), in which an assignment may map a variable to a single value or a sum value (modeled as Linkean plurals). The latter is known as a referential plurality. The present paper works with PCDRT-type assignments, as stated in Definition 2, and hence allow both singulars and plurals in the range of assignments.

Definition 2 (Assignment)
An assignment maps a variable to a possibly plural individual.

Introducing referential pluralities comes with a price tag, but it is necessary to deal with domains that lack well-defined atomicity, like the domains of mass nouns, atelic events, and spatial and temporal intervals (see, for example Bach 1986). In this paper, referential pluralities are needed to handle measure phrases like *two pounds of chicken* or *six gallons of water*, which may not introduce atomic parts. The system thus developed has two types of pluralities. *Discourse plurality*, also known as evaluation plurality, comes from considering a plurality of assignments. *Referential plurality*, also known as domain plurality, comes from considering a single assignment function that assigns a plurality to a variable.

Now that we have laid out the the architectural design of an info-state, we are almost ready to discuss how various types of updates work in this framework. Before we do that, let me define operations for extracting parts of an info-state. An info-state G with a set of assignments can be restricted to a sub-info-state with a set of assignments satisfying certain requirements. For example, (40) delivers a sub-info-state with all the assignments associating x with a. (41) delivers a sub-info-state with all the assignments associating x with some value in A.

Definition 3 (Sub-info-states)
(40) $G|_{x=a} := \{g \mid g\,x = a\}$

(41) $G|_{x \in A} := \{g \mid g\,x \in A\}$

In addition to restricting the assignments (i.e., the rows), we can also restrict the variables (i.e., the columns). We use the notion in (42) to project (or extract) all the values associated with the variable x in an info-state. Combining sub-info-states and value projection yields a mode of value projection relative to sub-info-states, as exemplified in (43) and (44).

Definition 4 (Value projections)

(42) $G(x) := \{g\,x \mid g \in G\}$

(43) $G|_{x=a}(y) := \{g\,y \mid g \in G\ \&\ g\,x = a\}$

(44) $G|_{x \in X}(y) := \{g\,y \mid g \in G\ \&\ g\,x \in X\}$

The first type of update to be discussed is variable introduction, also known as random assignment. It is contributed by existential quantification, as assumed in standard dynamic semantics. What variable introduction does is introduce all values in a set D as values for a new variable x, as defined below:

Definition 5 (Variable introduction/random assignment)

$G[\![\exists x]\!]H = \mathbb{T}$ iff there is a set D such that $H = \{g^{x \to a} \mid g \in G\ \&\ a \in D\}$, where a may be atomic or plural individuals.

This version of variable introduction falls in between the the corresponding operations in PDlL and PCDRT, in the sense that it may introduce referential pluralities but is still dependency-free.[29] These design features are empirically motivated. As Criterion 1 discussed at the beginning of this section reveals, a suitable logic for handling the monotonicity constraint needs to model pluralities and measure phrases. Given that measure phrases like *two pounds of chicken* and *six gallons of water* may involve measurement of objects that lack well-defined atomicity, it is hence assumed that a single assignment in this logic is free to assign any value, atomic or plural, to a variable. In this respect, variable introduction aligns with PCDRT rather than DPlL. However, departing from PCDRT but siding with DPlL, variable introduction in the present logic does not result in *variable dependence* between the introduced variable and any existing variable.[30] The formal definition of variable dependence is given below:

Definition 6 (Variable dependence)

In an info-state G, a variable y is dependent on a variable x if and only if there are a, b such that $a \neq b$ and $G|_{x=a}\,y \neq G|_{x=b}\,y$.   (see also Nouwen 2003:84)

In plain words, when the values associated with y does not remain constant relative to different values of x, then y is dependent on x. We may also say that the assignment of values to y co-varies with the assignment of values to x. Inhibiting variable dependence when a new variable

---

[29]Whether or not assignment functions may range over pluralities and whether or not new variable introduction is dependency-free are two independent design choices of a dynamic plural logic. PCDRT departs from DPlL in both choice points but the present logic only departs from DPlL in the first choice point. In this regard, the present logic is closer to DPlL.

[30]This is not quite true for the version of PCDRT adopted in Henderson (2014), which requires variables to have only one value at the evaluation level by default. This is an interesting strategy to remove dependencies introduced into discourse by the powerful variable introduction mechanism in PCDRT. The variable introduction defined here can be seen as a more 'automatic' way of getting rid of undesirable dependencies—they are not generated in the first place.

is introduced is a means to not generate spurious dependencies (van den Berg (1996):133–135). It forces variable dependence to come from a restricted source, such as the use of a distributive operator. I will return to this point shortly.

An example should help us see how variable introduction in the present framework works. Assume a model with a domain D containing the individuals c, d and c⊕d. Introducing a variable x with $\exists x$ into an info-state G involves the steps in (45). A visualization of this process can be found in Figure 8.

(45) Steps for introducing a new variable x to an info-state G
   a. Pick a non-empty subset D of values from D_e
   b. For each g in G and each value d in D, extend g to include x in the domain of g and d in the range of g, and collect the results in a set of assignments H (i.e., an info-state). The number of assignments in H is the cardinality of the cross product of G and D ($|G \times D|$). So, if G has two assignments and D is a singleton, there are two assignments in H. If G has two assignments and D has two members, then H has four assignments, so on and so forth.
   c. Repeat the above steps for each non-empty subset D of D_e. The total number of output info-states generable from introducing x into G is the cardinality of the power set of D_e minus the empty set (i.e., $2^{|D_e|} - 1$). So, if D_e has three members, then the total number of output info-states are seven.

H_1	y	x
h_1	a	c
h_2	b	c

H_2	y	x
h_1	a	d
h_2	b	d

H_3	y	x
h_1	a	c⊕d
h_2	b	c⊕d

H_4	y	x
h_1	a	c
h_2	b	c
h_3	a	d
h_4	b	d

G	y
g_1	a
g_2	b

H_7	y	x
h_1	a	c
h_2	b	c
h_3	a	d
h_4	b	d
h_5	a	c⊕d
h_6	b	c⊕d

H_6	y	x
h_1	a	d
h_2	b	d
h_3	a	c⊕d
h_4	b	c⊕d

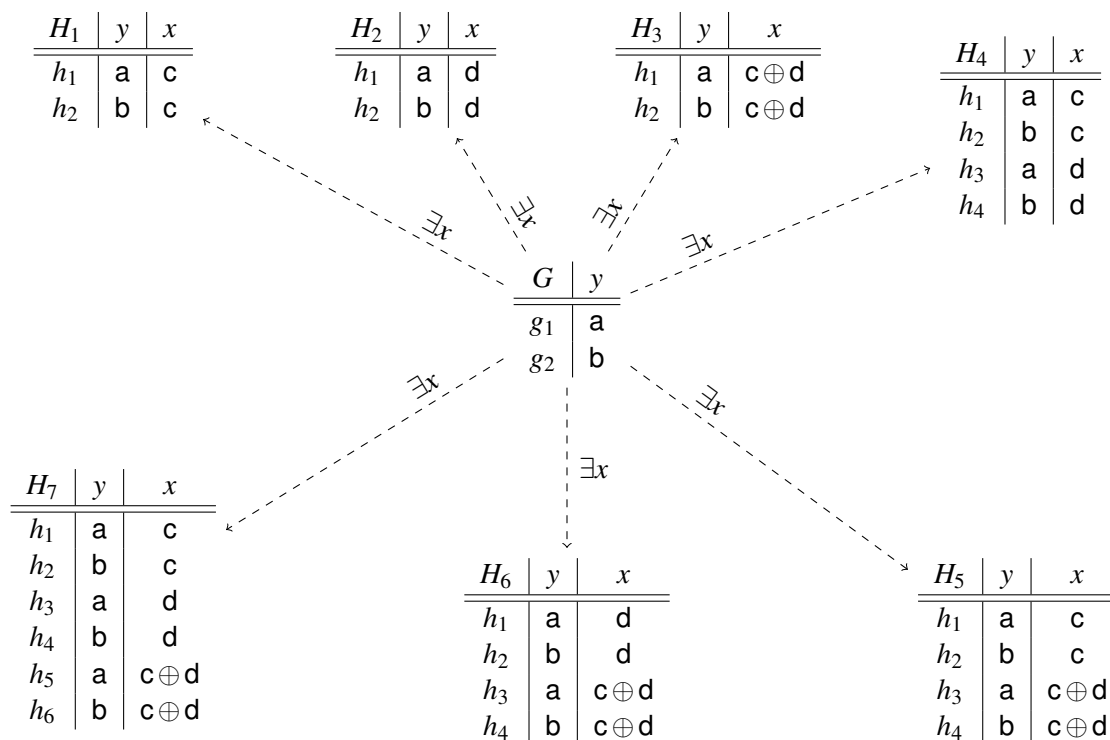| H_5	y	x
h_1	a	c
h_2	b	c
h_3	a	c⊕d
h_4	b	c⊕d

Figure 8: Variable introduction

It should now be clear that in none of the output info-states in Figure 8 is there variable dependence between the newly introduced variable x and the extant variable y. For variable dependence to be possible, a distributive operator is needed to scope over the operation of variable introduction. We turn to the distributive operator next.

To induce distributive, a distributive operator is used. The definition of the distributive operator is given below:

Definition 7 (Distributivity)

$$G[\![\delta_x(\phi)]\!]H = \mathbb{T} \text{ iff } G\,x\, = H\,x\, \;\&\; \forall \alpha \in G\,x.G|_{x=\alpha}[\![\phi]\!]H|_{x=\alpha} = \mathbb{T}$$

The distributive operator splits up the input info-state into substates based on the values stored in the subscripted variable. It then checks that the formula in its scope, i.e., ϕ, holds for each sub-state. Hence, for each sub-state, a distributivity update generates a set of output sub-states. These sets of sub-states are then pointwisely put back together to form the output info-state. If ϕ carries with it an existential quantifier, the new variable gets passed to the output. This way, a door is opened up for introducing variable dependence into info-states.

Let's take a concrete example to see how δ_x works. Suppose we have a formula with an existential expression in the scope of a distributive operator, i.e., $\delta_x(\exists y)$. Regarding the info-state G, $\delta_x(\exists y)$ first splits up the input info-state along the x dimension, resulting in two atomic sub-states, as shown in Figure 9 (on page 23).
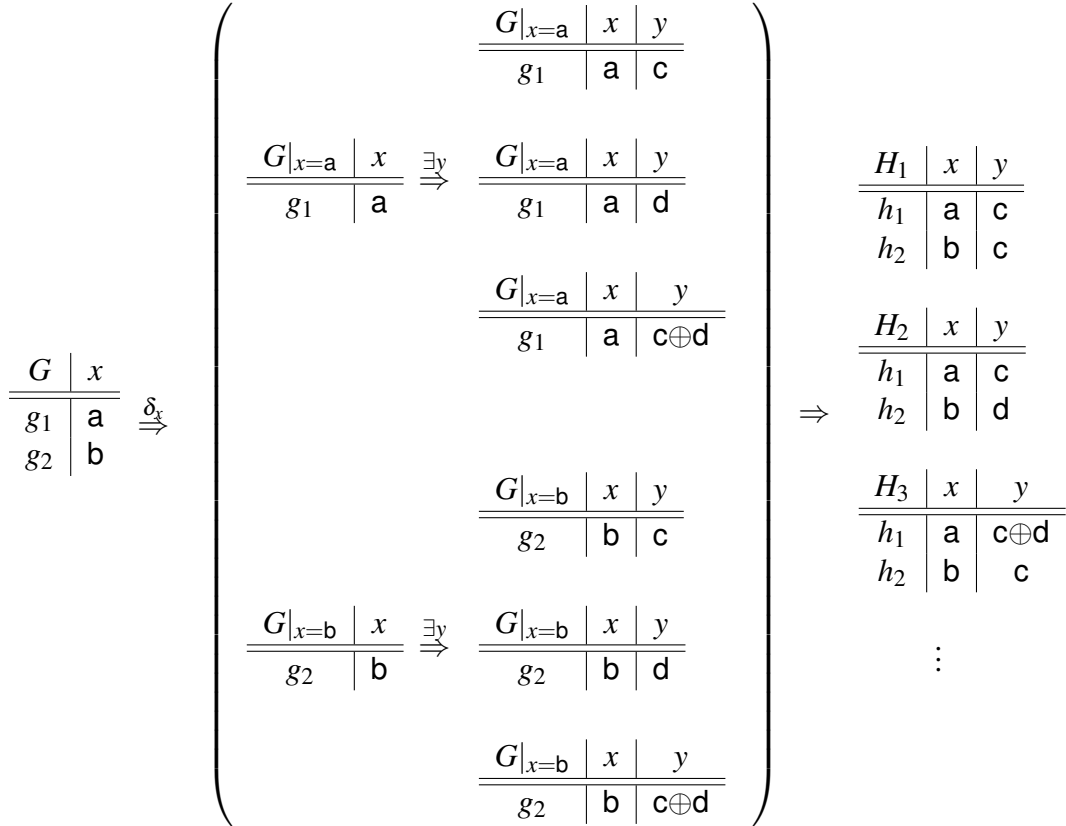

Figure 9: Introducing variable dependence with help of a distributive operator

Then intermediate sub-states are created by updating y to each of the two atomic sub-states and assigning random values to y. Note that within each 'leg' of the distributive update, there is no dependence relation between x and y. However, by collecting an intermediate output from each leg of the update, a set of global output info-states can be formed and variable dependence between x and y can be found in some of these global output states. For example H_2 and H_3 in this case.[31]

---

[31]The possible output info-states after evaluating distributivity in Figure 9 is the cross product of the two sets

For concreteness, as shown in (46-a) and (46-b), y stores different values in H_2 when x is restricted to different values. So, In H_2, y is dependent on x (and vice versa), according to the Definition 6 (Variable dependence).

(46)   Variable dependence in H_2 of Figure 9

    a. $H_2|_{x=\mathsf{a}}\, y = \{\mathsf{c}\}$

    b. $H_2|_{x=\mathsf{b}}\, y = \{\mathsf{d}\}$

4.1.2   Lexical relations, measurement, and logical connectives

Next, we turn to lexical relations, which only test input info-states but not modify them. Siding with DPlL but not PCDRT, lexical relations are satisfied *collectively*, as instructed by the following definitions:[32]

Definition 8 (Lexical relations)

(47) $G[\![R\, x_n, ..., x_1]\!]H = \mathbb{T}$ iff $G = H$ & $\left\langle \bigoplus G(x_1), ..., \bigoplus G(x_n) \right\rangle \in I(R)$

(48) $G[\![x =_\oplus a]\!]H = \mathbb{T}$ iff $G = H$ & $\bigoplus G\, x\ = a$, where x is a variable and a an individual.

The summation operator in the above definitions are used to bring a set of individuals (i.e., a discourse plurality) to a sum individual (i.e., a referential plurality) when checking for lexical relations. The definition of this operation is given in Definition 9. In more detail, (47) says that a sequence of n variables satisfies the lexical relation R if and only if the n-tuple arising from collecting all the values of each variable in the sequence belongs in the denotation of R. (48) is a special lexical relation that holds between a variable and a constant when the collective value assigned to the variable by a set of assignment functions is the same as the value the individual in the model. This piece of notation is used when a variable is set to a fixed value, such as a proper name.

Definition 9 (Discourse-level summation)

$\bigoplus G\, u = \bigoplus \{g\, u\ : g \in G\}$

Measurement is a special form of lexical relations. For this reason, just like lexical relations, measurement is also collectively evaluated. In particular, it is done by collapsing all evaluation-level pluralities into a domain-level plurality and applying a measure function to it.[33] The measure function comes with a parameterized dimension (such as height or volume), which

---

of outputs obtained from splitting the evaluation. Since there are three info-states in each leg of the evaluation, the total number of info-states in the final output is nine. Only three sample info-states are shown for space reasons.

[32]In Brasoveanu's (2008) PCDRT, lexical relations are *distributively* checked. This is a consequence of the dependency-introducing variable introduction in PCDRT. Not going into too much detail, we can understand the connection between variable introduction and lexical relations in the following way: since variable introduction makes available dependencies, collectively evaluating lexical relations amounts to not using those dependencies. As a result, the dependencies brought about by variable introduction are wasted. To effective use dependencies arising from variable introduction, a lexical relation should be distributively checked.

[33]I'm using the same symbol to stand for an object language measure function and a meta-language measure function here. The measure function can be decomposed into two parts: a function that relates entities to degrees and a function that relates degrees to numbers (Lønning 1987; Champollion 2017). For example, suppose John weights 68 kilograms. The measure function u_{weight} maps John to a degree, and then the unit function kilogram maps the degree to the number 68.

is omitted when irrelevant.[34] The definition for (generalized) measurement is given in (49). Cardinality measurement is taken to be a special form of measurement, one that requires the object being measured to have atomic parts, as reflected by (50).

Definition 10 (Measurement)

(49) $G[\![\mu_{\text{dimension}}\, x = d]\!]H = \mathbb{T}$ iff $G = H$ & $\mu_{\text{dimension}}(\bigoplus H\, x) = d$

(50) $G[\![\mu_{\text{card}}\, x = d]\!]H = \mathbb{T}$ iff $G = H$ & $|\{y \mid y \le \bigoplus H(x)\ \&\ y \text{ is an atom}\}| = d$

Finally, logical connectives are interpreted in accordance to DPL. The only logical connective used in this paper is dynamic conjunction. It is defined as follows.

Definition 11 (Dynamic conjunction)

$G[\![\phi \wedge \psi]\!]H = \mathbb{T}$ iff there is $K. G[\![\phi]\!]K = \mathbb{T}$ & $K[\![\psi]\!]H = \mathbb{T}$

As standardly assumed, conjunction is both internally and externally dynamic. Since it is internally dynamic, if variables are introduced in the first conjunct, their values and dependencies are available for the interpretation of the second conjunct (though not vice versa). Since it is externally dynamic, values and dependencies associated with variables introduced in both conjuncts are available outside the scope of conjunction.

4.2   Monotonicity constraint

Recall that in section 3, I have sketched the main proposal of this paper: binominal *each* introduces a monotonic measurement condition. The condition checks the the monotonic property of a measure function relative to the dependency established by distributive quantification. The definition, originally given in (27), is repeated in (51) below.

(51)   **Monotonic measurement in association with distributivity (with f)**
        A measure function μ is monotonic relative to a dependency f iff both of the following conditions hold:

   a.   Non-constant mapping
        There are distinct $b, b' \in \text{Dom}\, f.\ \mu(f\, b) \neq \mu(f\, b')$
   b.   Non-decreasing mapping
        For all $a, a' \in \text{Dom}\, f.\ a \le a' \rightarrow \mu(fa) \leqslant \mu(f\, a')$, and

The function f maps values stored in the distributivity key to corresponding values stored in the host. Since DPlL is chosen to couch the monotonicity condition because it represents dependencies among variables, it is natural that we can find a correlate for f in DPlL. The natural correlate of f in DPlL is an info-state storing the values contributed by the Key (i.e., $\text{Dom}(f)$) and the values contributed by the host (i.e., $\text{Ran}(f)$). Since distributive quantification is dynamic, not only can we retrieve values associated with the Key and the host of binominal *each*, we can also make reference to the dependency between the Key and the host. More importantly, since an info-state is a plural entity, i.e., a set of assignments, it has a natural mereological structure so we can easily talk about the mereological structure of a dependency, which is crucial for checking the monotonicity condition.

Given this background, we can now translate the monotonicity condition in (51) as a form of test in DPlL, as done below:

---

[34]The dimension parameter on a measure function determines which attribute of an object is being measured. For example, $\mu_{\text{height}}(\text{john})$ measures John's height while $\mu_{\text{weight}}(\text{john})$ measures John's weight. The dimension parameter is usually taken to be contextually determined (see Schwarzschild 2006, Wellwood 2015).

Definition 12 (Monotonic measurement in association with distributivity (DPlL))

(52) $G[\![\mathsf{dm}_{x,y}\ \mu]\!]H = \mathbb{T}$ iff

     a. $H = G$

     b.   There are distinct $B, B' \subseteq G(x)$. $\mu\left(\bigoplus G|_{x \in B}\ y\right) \neq \mu\left(\bigoplus G|_{x \in B'}\ y\right)$

     c.   For all $A, A' \subseteq G\ x.\ A \subseteq A' \rightarrow \mu\left(\bigoplus G|_{x \in A}\ y\right) \leqslant \mu\left(\bigoplus G|_{x \in A'}\ y\right)$

Let me link the various pieces in Definition 12 to the monotonicity condition in (51). The monotonicity condition is referred to as 'dm' (with d a mnemonic for distributivity and m a mnemonic for monotonicity and measurement. The measure function in Definition 12 corresponds to the same measure function in (51). As will be illustrated in the next subsection, it is contributed by a counting quantifier or a measure phrase. f in (51) is essentially G in Definition 12 as restricted to two x and y: x stores the values associated with the Key, i.e., $\mathsf{Dom}(f)$ and y stores values associated with the host, i.e., $\mathsf{Ran}(f)$.

To say that dm bears anaphoric indices amounts to attributing anaphoricity to binominal *each*. It is not a novel assumption, as there is a longstanding tradition in granting binominal *each* an anaphoric component. The tradition started in the early work of Burzio (1986) and Safir and Stowell (1988) and was later adopted in Dotlačil (2012) and Kuhn (2017).

To check whether a measure function satisfies dm, we need to access the values stored in the variable the measure function applies to. Definition 12 says that the measure function μ is monotonic on the dependency between x and y iff

- (52-a): Checking dm does not change the info-state in any way (i.e., it's a test).

- (52-b): In the input info-state, there are at least two sub-parts storing different x''s values that are also associated with different measurements of y's values.

- (52-c): Measuring y's values in an info-state storing less x's values does not yield a bigger number (or degree) than measuring y's values in an info-state storing more x's values.

Note that x and y may happen to be *the same* variable, as will be argued to be the case with predicative measure phrases like *two pounds*, which describe the measurement of entities introduced outside of the measure phrase. When this happens, a special version of the monotonicity condition, i.e., $dm_{x,x}$, results:

(53) $G[\![\mathsf{dm}_{x,x}\ \mu]\!]H = \mathbb{T}$ iff

     a. $H = G$

     b.   There are distinct $B, B' \subseteq G(x)$. $\mu\left(\bigoplus G|_{x \in B}\ x\right) \neq \mu\left(\bigoplus G|_{x \in B'}\ x\right)$

     c.   For all $A, A' \subseteq G\ x.\ A \subseteq A' \rightarrow \mu\left(\bigoplus G|_{x \in A}\ x\right) \leqslant \mu\left(\bigoplus G|_{x \in A'}\ x\right)$

In addition, I propose that the monotonicity condition in (53) is introduced as an **output context constraint** in the sense of Farkas (2002b) and Lauer (2009, 2012). In particular, (53) is treated as a constraint that is checked after the at-issue content has been established. If the at-issue content cannot pass the test, then the truth condition denoted by the sentence is not defined.[35] As a result, the sentence is *undefined*, rather than *false*. This is to model the fact that

---

[35] As summarized in Kuhn (to appear), there are a few types of 'postsuppositional' meanings (implemented in various ways), i.e., meanings introduced by expressions within a sentence (or a clause) but are delayed in evaluation until the sentence has been interpreted to give rise to an output context. The particular type of meaning taken up here belongs to what he calls 'blind post-suppositions', i.e., delayed meanings that must be satisfied within a local clause and returns ungrammaticality otherwise.

	ϕ	$\mathbb{T}$	$\mathbb{F}$
ψ			
$\mathbb{T}$		$\mathbb{T}$	$\mathbb{F}$
$\mathbb{F}$		$\star$	$\star$

Table 2: Constraint satisfaction

sentences with binominal *each* that fail dm (for various reasons) are judged unacceptable and not false, as illustrated below:

(54) a. *The drinks are 60 degrees (Fahrenheit) each.
      b. *The boys read some books each.
      c. *The boys read one book each, namely *Emma*.

An output constraint is interpreted following Definition 13. The special connective $\overline{\wedge}$ indicates that the constraint ψ applies after evaluating the at-issue content ϕ.

Definition 13 (Output context constraint)
$G[\![\phi \mathbin{\overline{\wedge}} \psi]\!]H = G[\![\phi]\!]H$ if $H[\![\psi]\!]H = \mathbb{T}$; otherwise, undefined.[36]

This definition says: the at-issue content given by ϕ, which represents the distributive component, has a truth value only if the output context of ϕ, which represents the monotonicity condition, admits ψ. A constraint behaves in a similar way to a presupposition in being a definedness condition, but it differs from a presupposition as the definedness condition is imposed on the *output* context, instead of the *input* context.

The determination of the truth condition of the entire formula $\phi \mathbin{\overline{\wedge}} \psi$ is given in Table 2 for concreteness. The top header ϕ represents the distributivity component and the side header ψ represents the monotonicity component. Let us first consider the predicted truth values for the entire sentence when the distributive component ϕ is true. According to the table, the entire sentence with an output context constraint is judged true if the constraint is satisfied but undefined if the constraint is not satisfied. This is in line with the data reported earlier in section 2—violation of the requirements subsumed by the monotonicity constraint leads to ungrammaticality (modeled as infelicity) but not falsity.

The situation is less ideal when the distributive component happens to be false. On one hand, the formulation correctly predicts that falsity of the distributive component also leads to falsity of the entire sentence even if the monotonicity condition is satisfied.[37] However, on the other hand, it *incorrectly* predicts that a sentence is *undefined* when both the distributive component and the monotonicity constraint are false. This is problematic as most speakers simply judge sentences with binominal *each* as false as long as the distributive component is false (see Table 5 in section 6). In other words, they no longer care about the satisfaction of the monotonicity constraint. This issue is taken up in more detail in section 6.

---

[36]As pointed out by a reviewer, this undefinedness is stated for a particular input-output pair. A sentence may be undefined for a particular input-output pair but defined for another pair. Given a particular input info-state, a sentence is undefined iff evaluating the sentence returns no defined output info-state.

[37]The monotonicity constraint can be accidentally satisfied in an info-state with prior values stored in the relevant variables, a design feature inherited from DPL and DPlL.

27

4.3 Composition

Although the original DPlL does not have sub-sentential compositionality, just like DPL, Nouwen (2003) and Brasoveanu (2008) have shown that it is possible to build sub-sentential compositionality into DPlL, following Muskens's (1996) compositional treatment of Discourse Representation Theory (Kamp 1981). In this section, I flesh out the compositionality assumptions adopted in the present framework.

To begin with, DPlL is treated as a typed logic. It includes basic types as listed in (55) and derived types $\tau \to \tau$ for functions.

(55) $\tau ::= e \mid t \mid s \mid d \mid \tau \to \tau$
     a. e: entities
     b. t: truth values
     c. s: (variable) assignments
     d. d: degrees

To keep type description reader-friendly, the following type abbreviations are used:

Name	Type	Abbr.	Variables	Examples
Info-state	$s \to t$	–	G, H	(table: x john, y sue; mary, peter)
proposition	$(s \to t) \to ((s \to t) \to t)$	t	ϕ, ψ	John left.
predicate	$e \to ((s \to t) \to ((s \to t) \to t))$	$e \to$ t	P, P'	pretty, book
quantifier	$(e \to$ t$) \to$ t	Q	$\mathscr{Q}$	every girl
measure functions	$e \to d$	m	m, m'	μ_{weight}

Table 3: Type abbreviations

I propose that all noun phrases hosting binominal *each* are measure phrases of some form. However, depending on whether a measure phrase occurs in an argument position, as in (56-a) and (56-b), or a predicate position, as in (56-c), it has slightly different types.

(56)   a.   John bought two apples.
     b.   John bought three pounds of chicken.
     c.   John is six feet (tall).

In an argument position, a measure phrase is a dynamic generalized quantifier (GQ), of type $(e \to$ t$) \to$ t. In a predicate position, a measure phrase is simply a predicate, of type $e \to$ t. However, unlike ordinary dynamic GQs and predicates, measure phrases have two additional components: a measure function and a measure head. The internal structures of different measure phrases are given in Figure 10.

Argumental measure phrases, analyzed as GQs, are shown in Figure 10a and Figure 10b. If the measure phrase is a cardinal GQ, the measure head is a silent determiner akin to the silent *many* in Hackl (2000). The measure head takes a degree (in the form of a number), a property and a measure function and returns a GQ. This measure head is defined in (57-a). If the measure phrase is a non-cardinal GQ, the measure head is assumed to be provided by a measure unit like *pound(s)*, which takes a degree, a property, and a measure function and returns a GQ, as defined in (57-b).
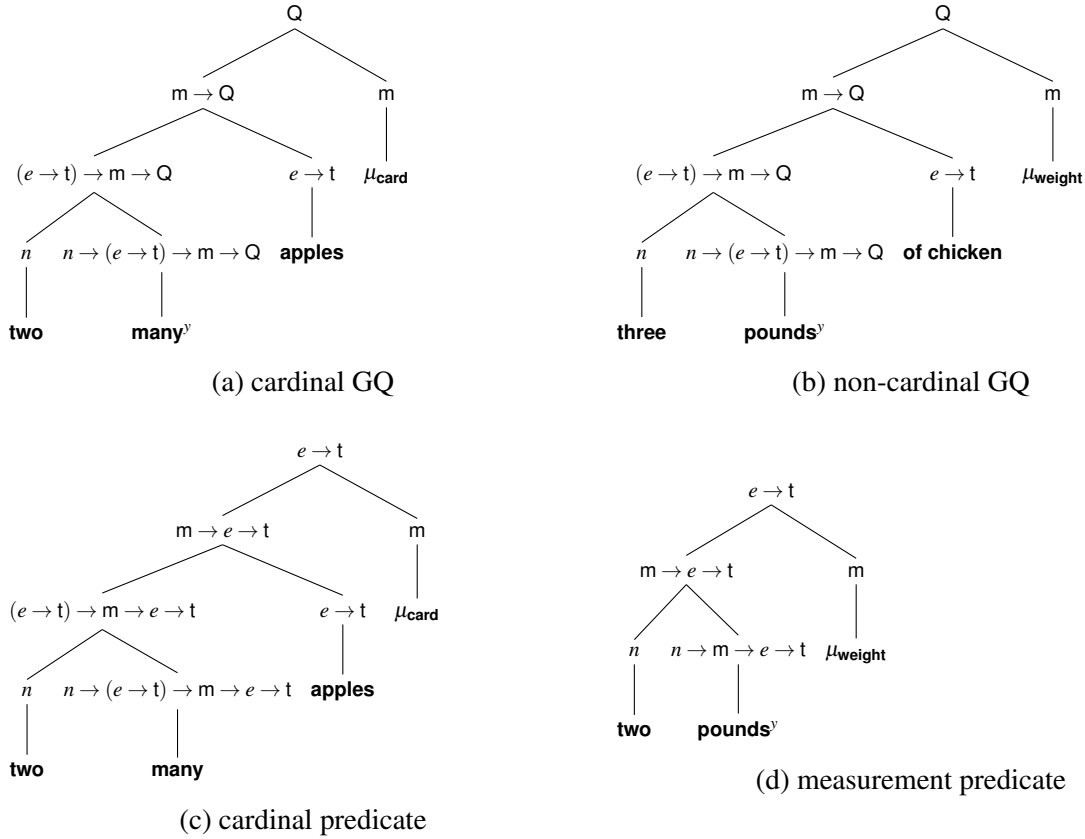
Figure 10: Argumental and predicative measure phrases

(a) cardinal GQ — tree with nodes: Q; $m \to Q$, m; $(e \to t) \to m \to Q$, $e \to t$, $\mu_{\textbf{card}}$; n, $n \to (e \to t) \to m \to Q$, **apples**; **two**, **many**^{y}

(b) non-cardinal GQ — tree with nodes: Q; $m \to Q$, m; $(e \to t) \to m \to Q$, $e \to t$, $\mu_{\textbf{weight}}$; n, $n \to (e \to t) \to m \to Q$, **of chicken**; **three**, **pounds**^{y}

(c) cardinal predicate — tree with nodes: $e \to t$; $m \to e \to t$, m; $(e \to t) \to m \to e \to t$, $e \to t$, $\mu_{\textbf{card}}$; n, $n \to (e \to t) \to m \to e \to t$, **apples**; **two**, **many**

(d) measurement predicate — tree with nodes: $e \to t$; $m \to e \to t$, m; n, $n \to m \to e \to t$, $\mu_{\textbf{weight}}$; **two**, **pounds**^{y}

(57) a. **many**$^y := \lambda n \lambda P \lambda m \lambda P'.\exists y \land P\, y \land P'\, y \land m\, y = n$
      b. **pound**$^y := \lambda n \lambda P \lambda m \lambda P'.\exists y \land P\, y \land P'\, y \land m\, y = n$ lbs

The cardinality measure head **many** selects (with help of agreement or some other means) a cardinality measure function $\mu_{\textbf{card}}$ (type $e \to n$), while a non-cardinality measure head like **pound** selects a non-cardinality measure function, like $\mu_{\textbf{weight}}$ (type $e \to d$). A measure function is assumed to be syntactically present and further away from a measure head, unlike that in Hackl (2000), which builds the measure function into the meaning of a measure head.

If a measure phrase is predicative and has a nominal predicate (e.g., *this is two pounds of chicken*), then the measure head takes the same ingredients, but returns a predicate rather than a GQ. The corresponding definitions of the predicative measure heads are given in (58-a) and (59-a). Lastly, sometimes a predicative measure phrase may not contain a common noun at all, as in *this is two pounds*. I assume that a measure head may optionally not take a nominal predicate as one of its arguments, giving rise to a simple, noun-less measure phrase. Sample definitions of the measure heads are given in (58-b) and (59-b).

(58) a. **many**$^{NP} := \lambda n \lambda P \lambda m \lambda u.P\, u \land m\, u = n$
      b. **many**$^{MP} := \lambda n \lambda m \lambda u.m\, u = n$

(59) a. **pound**$^{NP} := \lambda n \lambda P \lambda m \lambda u.P\, u \land m\, u = n$ lbs
      b. **pound**$^{MP} := \lambda n \lambda m \lambda u.m\, u = n$ lbs

With the assumptions about the internal structure of a measure phrase fleshed out, we are

now ready to add binominal *each* to the structure.[38]

I assume that binominal *each* attaches to a measure function and shifts the whole measure phrase to a higher-order meaning. Concretely, in a cardinal GQ, binominal *each* maps the GQ into a higher-order GQ by turning the measure function from an argument status (it is sought by a $m \to Q$ function) to a function status (it now seeks a $m \to Q$ function), as shown in Figure 11a. Similarly, in a measure phrase predicate, *each* attaches to the measure function and turns the whole measure phrase predicate into a higher order predicate, as shown in Figure 11b.

(a) a higher order dynamic GQ
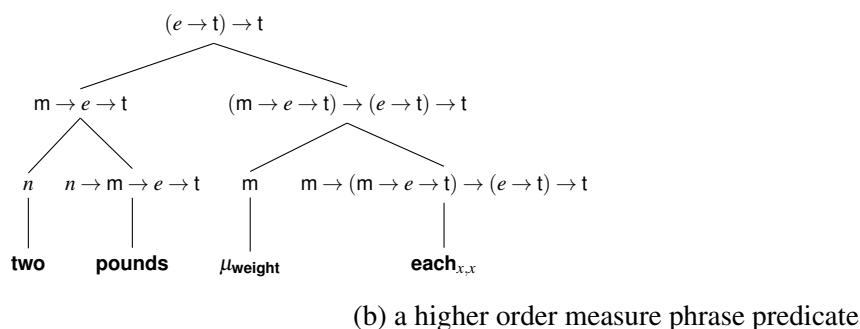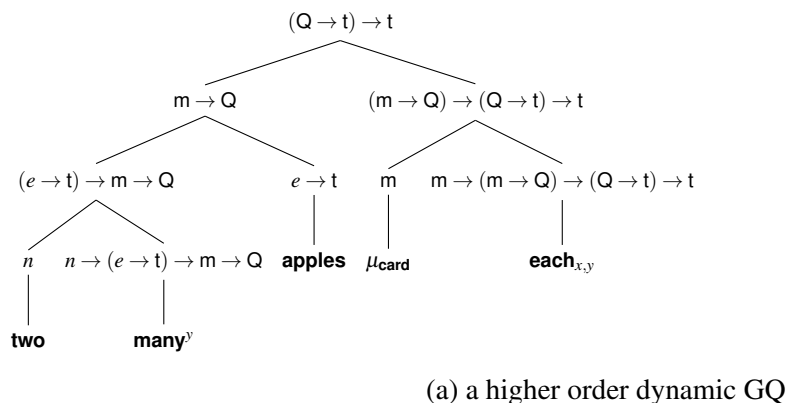

(b) a higher order measure phrase predicate

Figure 11: Binominal *each* gives rise to a higher-order meaning

Since binominal *each* can be hosted by both argumental and predicative measure phrases, and predicative measure phrases with or without a common noun component, we need to allow it to be type-polymorphic. I offer a schema for defining binominal *each* in (60-a), where f may range over any type α. In addition, when a measure phrase does not introduce any discourse

---

[38]As pointed out by a reviewer, not all argumental measure phrases contain a common noun other than the measurement unit. For example, *two miles* is an optional argument of the verb in (i) without a head noun. The two proposed lexical entries for binominal *each* cannot handle this use of measure phrases. The 'two-index' *each* for argumental measure phrases cannot be used because there is no individual d-ref introduced by the measure phrase that can be checked for monotonicity. The 'same-index' *each* for predicative measure phrases cannot be used because the measure function introduced by the measure phrase measures distance (of the walking events) but not the distance of the distributive key (i.e., the girls).

(i)   The girls walked two miles each.

An extension of the present framework has to be made to accommodate data like (i). A more straightforward extension involves the use of event d-refs (see Bittner (2014), Henderson (2014)), which can be used for in places of individual variables. In a less straightforward route, it may be assumed that measure phrases introduce degree d-refs that can be used for checking monotonicity, as done in Law (2019). The implications of incorporating degree d-refs into a plural logic framework remains to be precisified but it is worth noting that degree d-refs bear striking resemblance to the notion of degree pluralities in Dotlačil and Nouwen 2016 as both are basically collections of degrees.

variables, as in the case of a predicative measure phrase, *each* bears two instances of the the same anaphoric index, i.e., the anaphoric index for the variable storing the individuals measured by the measure function $\mu_{\mathbf{dim}}$. This is shown in (60-b).[39]

(60)   a.   **each**$_{x,y} := \lambda m \lambda f \lambda c. c(f\, m) \,\overline{\wedge}\, \mathsf{dm}_{x,y}\, m$ $\mathsf{m} \to (\mathsf{m} \to \alpha) \to ((\alpha \to \mathsf{t}) \to \mathsf{t})$

       b.   **each**$_{x,x} := \lambda m \lambda f \lambda c. c(f\, m) \,\overline{\wedge}\, \mathsf{dm}_{x,x}\, m$ $\mathsf{m} \to (\mathsf{m} \to \alpha) \to ((\alpha \to \mathsf{t}) \to \mathsf{t})$

As already can be seen in (60-a) and (60-b), after turning a GQ (or predicate) into a higher-order GQ (or a higher-order predicate), binominal *each* is capable of introducing a monotonic measurement constraint in a place different from where the original GQ (or predicate) takes scope. For example, in (60-a), the 'lower-order' GQ $f\,m$ takes scope inside c, but the monotonic measurement constraint is introduced outside c.

To see a concrete example, after composing with all the ingredients inside an argumental cardinal measure phrase, a host with binominal *each* essentially denotes a higher-order dynamic GQ, as shown in (61).[40]

(61)   **two many**^{y} **movies** $\mu_{\mathbf{card}}$ **each**$_{x,y}$ =

$$\lambda c. c\left(\lambda P. \exists y \,\wedge\, \mathsf{movie}\, y \,\wedge\, \mu_{\mathbf{card}}\, y = 2 \,\wedge\, P\, y \right) \overline{\wedge}\, \mathsf{dm}_{x,y}\, \mu_{\mathbf{card}}$$

This higher-order dynamic GQ looks for a function from GQ to truth values, puts the GQ (i.e., *two movies*) back in the scope of this function and introduces a monotonic measurement constraint outside the scope of this function (see Cresti (1995) and de Swart (2000) for higher-order quantifiers and Charlow (to appear) for their dynamic variants; Kuhn (to appear) is another study using higher-order dynamic GQs based on Charlow (to appear)). Figure 12 shows the Logical Form of a sentence with a higher-order dynamic GQ (the numerical indices induce a λ-abstraction rule, in the manner of Heim and Kratzer (1998)). This is essentially a 'split scope' mechanism that allows *two movies* to scope both inside and outside of distributivity. Scoping it inside distributivity gives us the correct narrow scope reading of *two movies* and scoping it outside of distributivity allows the monotonicity constraint to 'associate' with the internal structure of distributivity dependency.

Assuming the lexical entries in Table 4 for the definite NP *the girls*, the verb *saw* and the covert distributive operator, we obtain the final meaning of the LF, as shown in (62).

(62)   **two**^{y} **movies each**$_{x,y}$ $\underbrace{\left(\lambda \mathscr{Q}.\, \textbf{the girls}^x\ \textbf{dist}\ \underbrace{\left(\lambda u'. \mathscr{Q}\left(\lambda u.\ \textbf{saw}\ u\, u' \right) \right)}_{\beta} \right)}_{\alpha}$

---

[39] If we had taken binominal *each* to encode a distributive operator as part of its lexical meaning, the relevant lexical entry would need to be revised as follows:

(i)   **each**$_{x,y} := \lambda m \lambda f \lambda c. \delta_x(c(f\, m)) \,\overline{\wedge}\, \mathsf{dm}_{x,y}\, m$ $\mathsf{m} \to (\mathsf{m} \to \alpha) \to ((\alpha \to \mathsf{t}) \to \mathsf{t})$

To complement this entry, the plural subject has to take scope over binominal *each* and its host.

[40] All lexical predicates are cumulatively closed by default, following the assumptions in Landman (2000), Kratzer (2007), Brasoveanu (2013) and Champollion (2017). For example, if $x \in$ girl and $y \in$ girl, then $x \oplus y \in$ girl; and if $\langle x, y \rangle \in$ saw and $\langle x'y' \rangle \in$ saw, then $\langle x \oplus y,\ x' \oplus y' \rangle \in$ saw. For this reason, I do not mark a predicate with "*" to indicate cumulative closure.
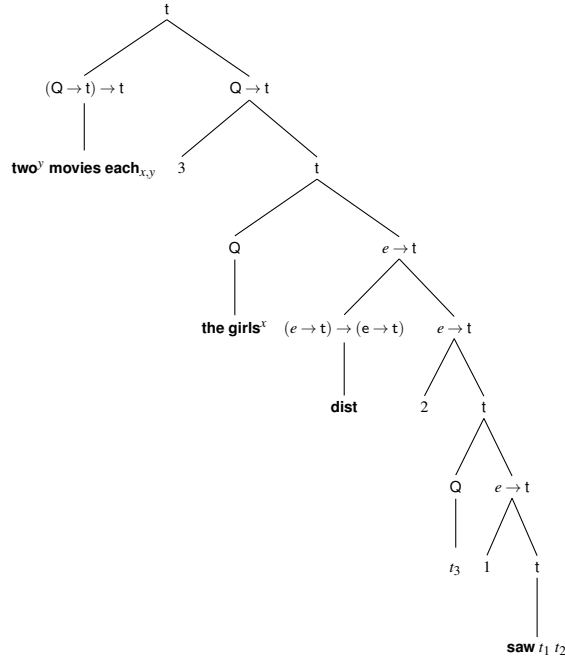
Figure 12: Scope taking of a higher order dynamic GQ

t
- $(Q \to t) \to t$: **two^y movies each**$_{x,y}$
- $Q \to t$
  - 3
  - t
    - Q : **the girls**^{x}
    - $e \to t$
      - $(e \to t) \to (e \to t)$: **dist**
      - $e \to t$
        - 2
        - t
          - Q : t_3
          - $e \to t$
            - 1
            - t : **saw** $t_1\ t_2$

Expression	Denotation	Type
the girls^x	$\lambda P.\mathsf{max}^x(\mathsf{girl}\ x) \wedge P(x)$	Q
saw	$\lambda u \lambda u'.\mathsf{saw}\ u\ u'$	$e \to e \to t$
dist	$\lambda P \lambda u.\delta_u(\mathsf{atom}\ u \wedge P\ u)$	$(e \to t) \to (e \to t)$

Table 4: Definite NPs, verbs and the distributive operator

a. $\quad \beta = \lambda \mathscr{Q}.\ \mathsf{max}^x\big(\mathsf{girls}\ x\big) \wedge \delta_x\Big(\mathsf{atom}\ x\ \wedge\ \mathscr{Q}\big(\lambda u.\ \mathsf{saw}\ u\ x\big)\Big)$

b. $\quad \alpha = \mathsf{max}^x\big(\mathsf{girls}\ x\big) \wedge \delta_x \left(\begin{array}{l} \mathsf{atom}\ x\ \wedge\ \exists y\ \wedge\ \mathsf{movie}\ y\ \wedge \\ \mu_{\mathsf{card}}\ y = 2\ \wedge\ \mathsf{saw}\ y\ x \end{array} \right) \bar{\wedge}\, \mathsf{dm}_{x,y}\ \mu_{\mathsf{card}}$

As shown in (62), the split scope mechanism allows *two movies* to scope inside the distributive operator but dm to scope outside the distributive operator. The 'association-with-distributivity' effect is clearly seen in the dm test in (62-b). The test bears an anaphoric index x, which is the same index borne by the distributive operator. This anaphoric index stores values based on which an info-state is split up into sub-states to check a formula distributively.

To test for dm, we first assemble the distributivity update. Assuming a scenario in which three girls each saw two different movies, the output of the distributivity update can be visualized in Figure 13.
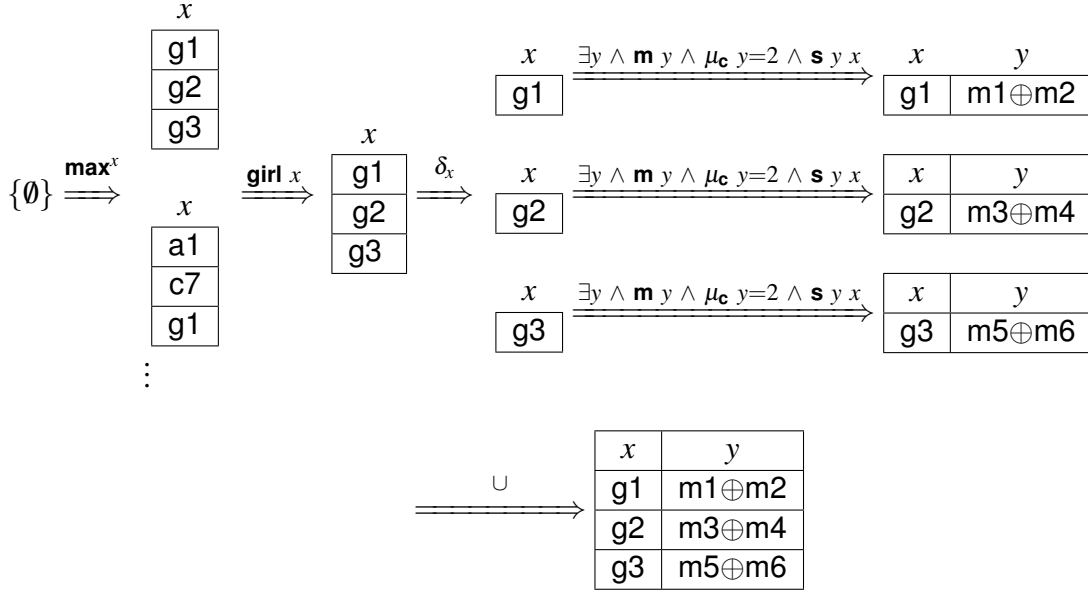
Figure 13: Distributivity update

The monotonic measurement constraint, spelled out in (63), is evaluated against the output of the distributivity update. It first requires that the info-state be split up into sub-states each storing one or more values in the variable x. With three values in x, seven such sub-states can be found (excluding the empty sub-state, which stores no value in x). Then, it compares these sub-states, requiring that if a sub-state whose x-value is a proper subset of the x-value of another sub-state, then measuring y's cardinality in the former sub-state does not yield a bigger number than measuring y in the latter sub-state.

(63) $G[\![\mathsf{dm}_{x,y}\ \mu_{\mathsf{card}}]\!]H = \mathbb{T}$ iff

    a. $H = G$ and

    b. $\exists B, B' \subseteq G\, x.\ \mu_{\mathsf{card}}\left(\bigoplus G|_{x \in B}\, y\right) \neq \mu_{\mathsf{card}}\left(\bigoplus G|_{x \in B'}\, y\right)$

    c. $\forall A, A' \subseteq G\, x.\ A \subseteq A' \rightarrow \mu_{\mathsf{card}}\left(\bigoplus G|_{x \in A}\, y\right) \leq \mu_{\mathsf{card}}\left(\bigoplus G|_{x \in A'}\, y\right)$

For concreteness, let's consider two info-states, shown in Figure 14, that verify dm. In info-state G, three girls each watched a different set of two movies. The cardinality of y (i.e., the movies) in each x sub-state is provided under the matrix. Since the cardinality of y never decreases in a bigger sub-state containing more x-values, non-decreasing mapping is satisfied. In addition, the cardinality of y is not constant in all the x sub-states, non-constant mapping is satisfied. As a result, dm is satisfied by Info-State G. Another info-state that also verifies dm is Info-State G', which has two girls seeing two identical movies but a third girl seeing two different movies. Again, this info-state satisfies both non-decreasing mapping and non-constant mapping, hence also dm.

    Of course, not all distributivity updates satisfy dm. If the values stored in y do not vary across the distributivity dependency, as in Info-State G''. in Figure 15, dm is violated. Recall that since dm is modeled as a constraint, the predicted judgment for the corresponding sentence containing a binominal *each* is infelicitous, or unacceptable, rather than false. This is how dm captures the variation inference triggered by binominal *each*.

    When the measure phrase is predicative, as in (64-a) and (64-b), the measure phrase does

G	x	y
g_1	g1	m1 $\oplus$ m2
g_2	g2	m3 $\oplus$ m4
g_3	g3	m5 $\oplus$ m6

Info-state G

$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g1}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g2}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g3}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g1, g2}\}}\, y) = 4$
$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g1, g3}\}}\, y) = 4$
$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g2, g3}\}}\, y) = 4$
$\mu_{\textbf{card}}(\bigoplus G|_{x\in\{\text{g1, g2, g3}\}}\, y) = 6$

G'	x	y
g'_1	g1	m1 $\oplus$ m2
g'_2	g2	m1 $\oplus$ m2
g'_3	g3	m3 $\oplus$ m4

Info-state G'

$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g1}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g2}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g3}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g1, g2}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g1, g3}\}}\, y) = 4$
$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g2, g3}\}}\, y) = 4$
$\mu_{\textbf{card}}(\bigoplus G'|_{x\in\{\text{g1, g2, g3}\}}\, y) = 4$

Figure 14: Info-states that verify *the girls saw two movies each*

G''	x	y
g_1	g1	m1 $\oplus$ m2
g_2	g2	m1 $\oplus$ m2
g_3	g3	m1 $\oplus$ m2

Info-State G''

$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g1}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g2}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g3}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g1, g2}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g1, g3}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g2, g3}\}}\, y) = 2$
$\mu_{\textbf{card}}(\bigoplus G''|_{x\in\{\text{g1, g2, g3}\}}\, y) = 2$

Figure 15: An info-state that fails to verify *the girls saw two movies each*

not introduce a discourse variable. dm is checked by just using a single discourse variable, i.e., the variable storing the values for the distributivity key (the relevant angles for (64-a) and the relevant coffees for (64-b)).

(64) a. The angles are 60 degrees each.
   b. *The coffees are 60 degrees each.

For concreteness, the interpretation of these two sentences are given below:[41]

---

[41] A reviewer correctly pointed out that the monotonicity condition contributed by the 'same-index *each*' does not require a distributive evaluation of a predicative measure phrase. To take (65-a) as an example, a collective evaluation of the predicative measure phrase yields the following interpretation (the distributive operator and the atomicity requirement are both removed):

(i) $\max^x\Big(\text{angles}\, x\Big) \wedge \big(\mu_{\text{angle}}\, x = 60°\big) \barwedge \text{dm}_{x,x}\, \mu_{\text{angle}}$

The reader can verify that since there is no novel d-ref y storing the measurement of each angle, the monotonicity requirement is easily satisfied assuming angle measurement is monotonic, regardless of whether each angle angle measurement is $60°$. As mentioned in footnote 16, this constitutes an argument for attributing distributivity to

H	x
h_1	a1
h_2	a2
h_3	a3

Info-State H

H''	x
h'_1	c1
h'_2	c2
h'_3	c3

Info-State H'

$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a1\}}\, x) = 60°$
$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a2\}}\, y) = 60°$
$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a3\}}\, y) = 60°$
$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a1,\, a2\}}\, y) = 120°$
$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a1,\, a3\}}\, y) = 120°$
$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a2,\, a3\}}\, y) = 120°$
$\mu_{\textbf{angle}}(\bigoplus H|_{x\in\{a1,\, a2,\, a3\}}\, y) = 180°$

$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c1\}}\, x) = 60°$
$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c2\}}\, x) = 60°$
$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c3\}}\, x) = 60°$
$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c1,\, c2\}}\, x) = 60°$
$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c1,\, c3\}}\, x) = 60°$
$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c2,\, c3\}}\, x) = 60°$
$\mu_{\textbf{temp}}(\bigoplus H'|_{x\in\{c1,\, c2,\, c3\}}\, x) = 60°$

Figure 16: Monotonic measurement (left) vs. non-monotonic measurement (right)

(65) a. $\mathsf{max}^x\Big(\mathsf{angles}\, x\Big) \wedge \delta_x\big(\mathsf{atom}\, x \,\wedge\, \mu_{\mathsf{angle}}\, x = 60°\big)\, \overline{\wedge}\, \mathsf{dm}_{x,x}\, \mu_{\mathsf{angle}}$

   b. $\mathsf{max}^x\Big(\mathsf{coffees}\, x\Big) \wedge \delta_x\big(\mathsf{atom}\, x \,\wedge\, \mu_{\mathsf{temp}}\, x = 60F\big)\, \overline{\wedge}\, \mathsf{dm}_{x,x}\, \mu_{\mathsf{temp}}$

The corresponding monotonic measurement constraints have a similar form, as shown in (66), differing only with respect to whether the values stored in x are angles or coffees, and whether the measure function measures angle degrees or temperature.

(66) $G[\![\mathsf{dm}_x\, \mu_{\mathsf{angle/temp}}]\!]H = \mathbb{T}$ iff
   a. $H = G$ and
   b. $\exists B, B' \subseteq G\, x.\ \mu_{\textbf{angle/temp}}\Big(\bigoplus G|_{x\in B}\, x\Big) \neq \mu_{\textbf{angle/temp}}\Big(\bigoplus G|_{x\in B'}\, x\Big)$
   c. $\forall A, A' \subseteq G\, x.\ A \subseteq A' \rightarrow \mu_{\textbf{angle/temp}}\Big(\bigoplus G|_{x\in A}\, x\Big) \leq \mu_{\textbf{angle/temp}}\Big(\bigoplus G|_{x\in A'}\, x\Big)$

As shown in the info-states in Figure 16, it is possible to satisfy dm if the measure function is monotonic, as in the case of $\mu_{\textbf{angle}}$ (Info-State H), but not if the measure function is non-monotonic, as in the case of $\mu_{\textbf{temp}}$ (Info-State H').[42]

As already pointed out in section 3, the monotonicity condition cannot be satisfied with a non-monotonic measure function even if the measure function associates different degrees for different values in the Key, as allowed in (67).

(67) *The coffees are more than 60 degrees each.

For concreteness, suppose interpreting (67) gives rise to the output info-state H'' in Figure 17. Since the three coffees all have different temperatures, H'' satisfies the non-constant mapping clause of the monotonic measurement condition. However, H'' fails the non-decreasing mapping clause, as it is not true that the sum of two coffees is always of a higher temperature than the individual coffees that make up the sum.

---

binominal *each*.

[42]When the measure function is non-monotonic, there is no way to satisfy dm. However, when the measure function is monotonic, whether or not dm is satisfied is context-dependent, as it matters what values are associated with the variable being measured, i.e., the host variable.

H''	x
h_1	c1
h_2	c2
h_3	c3

Info-State H''

$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c1}\}}\, x) = 60°$$
$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c2}\}}\, x) = 70°$$
$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c3}\}}\, x) = 80°$$
$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c1, c2}\}}\, x) = 65°$$
$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c1, c3}\}}\, x) = 70°$$
$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c2, c3}\}}\, x) = 75°$$
$$\mu_{\mathbf{angle}}(\bigoplus H|_{x \in \{\text{c1, c2, c3}\}}\, x) = 70°$$

Figure 17: Non-monotonic measurement giving rise to a non-constant mapping but not a non-decreasing mapping

4.4 Interim summary

I have demonstrated how to translate dm as an output constraint in DPlL$^{\oplus}$, a dynamic plural logic sharing properties with van den Berg's (1996) DPlL and Brasoveanu's (2008) PCDRT. The use of plural logic enables us to model distributivity-induced dependency as a discourse plurality, and marrying plural logic with a dynamic logic allows us to record this dependency and its internal structure. The anaphoric component on binominal *each* retrieves this dependency, and the monotonic measurement condition makes crucial use of the internal structure of this dependency.

5 Comparisons with and connections to previous studies

In this section, I compare the present approach to binominal *each* with previous studies on this marker.

5.1 Studies attributing the variation requirement to variable dependence

There are a few extant studies on dependent indefinites and markers like binominal *each* couched in PCDRT, a variant of DPlL (e.g., Henderson 2014, Champollion 2015, Kuhn 2017). Although there are some nontrivial differences between DPlL and PCDRT, they share the overall dynamic framework and the plural architecture. For this reason, the present account can be more or less directly compared with the accounts developed in these studies.

To begin with, Henderson (2014) investigates dependent indefinites in Kaqchikel and concludes that their variation requirement should be analyzed as imposing a post-suppositional plurality condition (known as *evaluation-level plurality*) on the functional dependency arising from distributive quantification. On the basis that binominal *each* also exhibits some form of variation requirement, Champollion (2015) and Kuhn (2017) take binominal *each* to be a dependent indefinite marker and extend Henderson's analysis of dependent indefinites to binominal *each* constructions. Although Henderson's analysis is modified in Kuhn (2017), the core of the analysis, namely, that dependent indefinites contribute an evaluation-level plurality condition, is shared by both Champollion (2015) and Kuhn (2017). Since the evaluation-level

plurality condition was devised to handle the variation requirement, the comparison focuses on its similarities to and differences from the monotonicity constraint.

Let me introduce the evaluation-level plurality condition with a concrete sentence like (68). This sentence can largely be translated into DPlL$^{\oplus}$ as in (69). The only bit that is not yet defined in our framework is the evaluation-level plurality condition in the last conjunct. Let me define it below, using Henderson's original definition.

(68)   The students hugged one dog each.

(69) $\mathsf{max}^x(\mathsf{student}\ x) \wedge \delta_x(\exists y \wedge \mu_{\mathsf{card}}\ y = 1 \wedge \mathsf{dog}\ y \wedge \mathsf{hug}\ y\ x) \wedge y > 1$

Definition 14 (Evaluation-level plurality (Henderson 2014))
$G[\![y > 1]\!]H = \mathbb{T}$ iff $G = H$ and $|\{h\ y\ |\ h \in H\}| > 1$

h in Definition 14 is a single assignment, so $h(y)$ yields a single value (which can be in the form of a plural individual). Since an info-state has a set of h-assignments ($h_1, h_2, ...h_n$), we can collect a set of $h(y)$ values ($\{h_1\ y, h_2\ y, ...h_n\ y\}$). An evaluational-level cardinality can be computed based on how many members are in this set. If all the assignments assign to y the same value, then there is only one member in the set. Such a set does not satisfy evaluation-level plurality. However, if at least two assignments assign different values to y, evaluation-level plurality is satisfied.

Since the evaluation-level plurality condition is evaluated after the distributivity quantification, each h that associates y with a value also associates x with a value, from the distributivity key. For this reason, requiring y to exhibit evaluation-level plurality following a distributive quantification has the same effect as requiring y to depend on x. To see this, consider the definition of variable dependence repeated from Definition 6.

Definition 6 (variable dependence)
In an info-state G, a variable y is dependent on a variable x if and only if there are a, b such that $a \neq b$ and $G|_{x=a}\ y \neq G|_{x=b}\ y$.   (see also Nouwen 2003:84)

If we set the variable x to store all the relevant values in the Key and y to store the values introduced by the host. The requirement that y is associated with different values for at least two distinct values in x is the same as saying that y is associated with at least two values at the evaluation level. So, we can safely conclude that evaluation-level plurality and value dependence are one and the same requirement.[43]

As already pointed out earlier, variable dependence is a prerequisite for the monotonicity constraint when counting quantifiers are concerned: a pair of (possibly identical) variables can satisfy the monotonicity constraint only if they exhibit variable dependence. For this reason, the monotonicity constraint is straightly stronger than variable dependence (or evaluation-level plurality): Variable dependence only handles the variation requirement while the monotonicity constraint also handles the counting quantifier requirement and the monotonic measurement requirement.

In addition to the primary difference between variable dependence and the monotonicity constraint, there are a few less pronounced differences between the present study and its predecessors. First, there is a difference in the kind of meaning status given to the evaluation-level plurality requirement and the monotonic measurement requirement. In Henderson (2014) and

---

[43]Evaluation-level plurality is notationally more economical as it only uses one variable, i.e., the variable introduced by a distributive numeral. The variable storing values contributed by the Key is not used explicitly. However, it is used implicitly, as y is in the distributive scope of x.

Champollion (2015), evaluation-level plurality is analyzed as a delayed at-issue test.[44] However, in both Kuhn (2017) and the present study, the corresponding component is analyzed as a not-at-issue meaning. The motivation for modeling it as a not-at-issue meaning is empirically driven—failure to satisfy the variation component leads to unacceptability rather than falsity. Although Kuhn (2017) calls his evaluation-level plurality constraint a 'presupposition' while I call the monotonicity constraint here an 'output constraint', the two essentially amount to the same thing. Kuhn (2017) calls the constraint a presupposition because it is placed on the *input* context for evaluating the constraint, which is precisely the *output* context of evaluating distributive quantification. I thank a reviewer for alerting me to this point.

Lastly, this study has adopted a higher order meaning approach to model delayed evaluation, following Charlow (to appear). By contrast, Henderson (2014) follows Brasoveanu (2013) and uses a post-supposition instead and the assumption carries over to Champollion (2015). Kuhn (2017) uses ordinary scope-taking without higher order meaning to model delayed evaluation. The merits and shortcomings of these strategies are discussed in the next subsection.

5.2 Split scope via other mechanisms

As briefly mentioned earlier, two other mechanisms have been explored to model split scope, namely, the use of post-suppositions and the use of scope-taking without higher-order meaning. In this subsection, I review these alternatives and why I opt for higher-order meaning.

Henderson (2014) is the first to use post-suppositions to model the variation requirement of dependent indefinites (see also Champollion 2015). However, Kuhn (2017)) points out that post-suppositions, without further assumptions, predict a lack of locality in the licensing of dependent indefinites. The prediction is not borne out, as binominal *each* and its distributivity key cannot be separated by a scope island. Consider the following examples (judgments due to the credited sources):

(70)   *The boys said Mary captured two snakes each.   (Safir and Stowell 1988:(48))

(71)   a.   Jones proved the prisoners guilty with one accusation each.
        b.   Bob made/let Sam and Tom leave on two occasions each.

(Safir and Stowell 1988:(36a-b))

(72)   ??The linguists thought two theories each were refuted.   (Simon Charlow, p.c.)

(73)   The linguists want two theories each to be refuted.   (Simon Charlow, p.c.)

In (70) and (72), binominal *each* and its hosts are inside tensed clauses, which have been independently identified as a scope island for quantifiers (e.g., May 1985, Beghelli 1995, Barker 2002, Charlow 2014). In (71) and (73), the dependent indefinites are inside ECM clauses, which have been observed not to be a scope island for quantifiers (e.g., May 1985). The fact that dependent indefinites introduced by binominal *each* are subject to the same locality conditions governing quantifier scope suggests that a locality-sensitive mechanism should be used for licensing dependent indefinites.

To model the island sensitivity of dependent indefinites, Kuhn (2017) suggests a scope-taking analysis, in which a noun phrase like *two theories each* has to undergo quantifier-raising (QR) to take wide scope.

A drawback of Kuhn's QR analysis (discussed in Kuhn 2017 and credited to an anonymous reviewer), is that it fails to account for the grammaticality of dependent indefinites with a bound

---

[44]This is not entirely correct as Henderson (2014) and Champollion (2015) have not made explicit what would happen when the condition fails.

pronoun inside them.

(74) Minden rendezö benevezte két-két filmjét.
       every   director entered   two-two film-POSS.-3SG-ACC
       'Every^x director entered two films of his$_x$ (in the competition).'

In this Hungarian example, the noun phrase restriction of the distributive numeral has a (possessor) pronominal bound by the quantifier that licenses the distributive numeral. If the distributive numeral has to take wide scope over its licensor to be licensed, then the pronoun is left unbound.

Based on considerations of island sensitivity and pronominal binding, Charlow (to appear) suggests a scope-taking mechanism involving higher order meaning. Charlow's higher-order meaning approach has a very similar empirical coverage as the post-supposition approach, with the exception of island sensitivity, which favors the former. In this study, I have adopted the higher-order meaning approach for it has better empirical coverage, although the choice is largely immaterial to the claim that binominal *each* makes reference the mereological structure of a distributivity dependency.

5.3 Studies in static semantics

There is a vast literature on binominal *each* couched in static semantics. It is beyond the scope of the present paper to offer a comprehensive review of previous studies on this topic. However, it is worth pointing out the major developments that have paved way for the ideas used in the present paper.

An early study on binominal *each* is Link (1987), which set the stage for treating binominal *each* as a distributive operator. This treatment is adopted in many subsequent studies, including Zimmermann (2002), Dotlačil (2012), Champollion (2010, 2017). However, these studies place their primary focus on the distributivity component and do not really recognize the variation component. As such, they differ quite drastically from the present paper, which takes the variation component as its primary concern.

There are a few studies that take up the variation component. For example, Safir and Stowell (1988) recognize a strong form of the variation inference, and conceive binominal *each* as a one-to-one distribution function, establishing a one-to-one correspondence between elements in the distributivity key to elements in the distributivity share. This strong form of variation is later criticized by Moltmann (1991) and Zimmermann (2002). Cable (2014) extends the semantics established for dependent indefinites in Tlingit to binominal *each*, arguing that *each* is both a distributive marker and bears a variation inference. Despite recognizing the variation component, these studies either fail to account for the counting quantifier requirement and/or the monotonic measurement requirement.

Despite these differences, studies in the static tradition have offered great insights to the study of binominal *each* in the present work. For one thing, it has been a longstanding puzzle how binominal *each* access the distributivity key. The received wisdom is that there are null pronouns in the NP that hosts binominal *each* that help connect it with the distributivity key, as suggested in Safir and Stowell (1988). This idea is further refined in Zimmermann (2002), with the pronoun treated as an anaphoric index directly borne by binominal *each*. The strategy is then imported into a dynamic framework by Dotlačil (2012) and adopted in Kuhn (2017) and the present work.[45]

---

[45]The anaphoric index provided by the distributivity key is not used in Henderson 2014, as his formulation of the evaluation-level plurality condition does not need direct reference to the distributivity key.

Many studies also share the intuition that *each* is a marker of quantificational dependence. Choe (1987) and Milačić et al. (2015) are notable examples. This intuition is also relevant in the present study, albeit in a slightly different manner. In previous studies, the core contribution of binominal *each* is to signal quantificational dependence. However, in the present study, the core contribution is a variation component formalized in terms of a monotonicity constraint. A separate constraint is needed because quantificational dependence is a *necessary but not sufficient* condition for using binominal *each*.

The present study is not the first to observe that monotonicity plays a role in conditioning the distribution binominal *each*. Based on essentially the same range of data reported in section 2, Zhang (2013) sketches an analysis that treats binominal *each* as an operator expressing a ratio over two measurements, which is similar to the contribution of the word *average* as suggested in Kennedy and Stanley (2009). The monotonicity requirement is taken to be a presupposition associated with the measurement constructions (see also Schwarzschild 2006, Wellwood 2015).

As a final note, the monotonicity constraint is related to the idea of 'structure-preserving binding', developed in Jackendoff (1996) to deal with a host of phenomena ranging from telicity to quantification. Jackendoff suggests to broaden the notion of binding from a relation between two identical variables to a relation between two variables that are linked in some way. Most importantly, he argues that it is fruitful to study the links in terms of structure-preserving maps. He implements structure-preserving binding in the framework of Conceptual Semantics, which differs from the framework used in this work substantially. However, the core of the idea of structure-preserving binding resonates with the notion of the monotonicity constraints developed here.

6 Remarks on output context constraints

6.1 Truth-value gap

In section 4.3, I suggest modeling the monotonic measurement condition of binominal *each* as an output context constraint, following Farkas (2002b), Lauer 2012 and Kuhn (2017). This is motivated by the presence of a truth value gap in the meaning of sentences with binominal *each*. In particular, when the distributivity component is judged true, the monotonicity component cannot typically lead to the falsity of the entire sentence with binominal *each*. As an example, the entire sentence in (75) is either judged as *true* (when the condition is satisfied, as in Figure 18a) or *unacceptable* (when the condition is not satisfied, as in Figure 18b) (see also Kuhn 2017).

(75)   The girls saw one movie each.

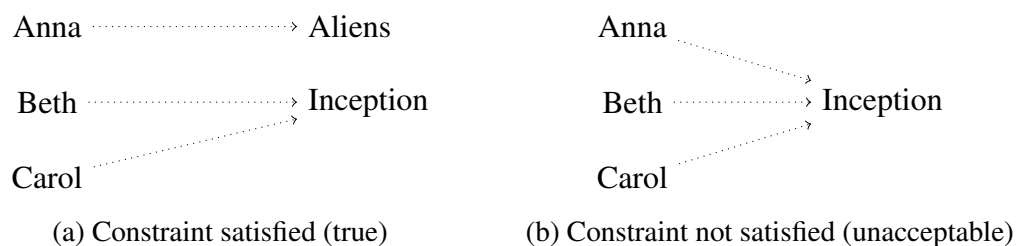(a) Constraint satisfied (true)   (b) Constraint not satisfied (unacceptable)

Figure 18: The truth value of a sentence containing binominal *each* when distributivity is satisfied

This kind of truth value gap is characteristic of input context constraints, i.e., presuppositions, and have been attributed to output context constraints (Farkas 2002b, Lauer 2012, Kuhn 2017, to appear).

A natural question that arises is what would happen if distributivity is *not* satisfied. This is particularly interesting because in DPlL distributivity has to be true to establish the dependency needed for checking the monotonicity constraint. When distributivity is false, there is no output with dependency that can support the constraint and hence the constraint cannot be satisfied. Since the constraint is a definedness condition, like a presupposition (Strawson 1950, Heim 1983), the prediction is that the falsity of the distributivity component would lead speakers to judge a sentence with binominal *each* as unacceptable. However, this prediction is not borne out.[46] As shown in the two situations in Figure 19, a sentence with binominal *each* is judged *false* rather than unacceptable or undefined when distributivity is *not* satisfied. This is regardless of whether a partial dependency may help satisfy the constraint, as in Figure 19a, or not, as in Figure 19b.

(a) Constraint satisfied by partial dependency (false)

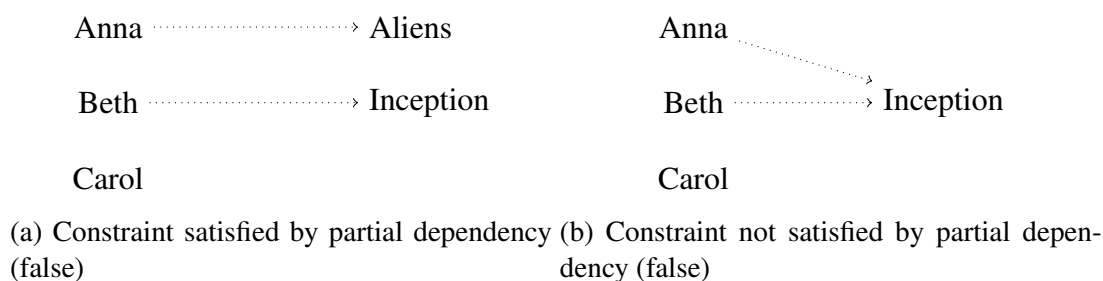(b) Constraint not satisfied by partial dependency (false)

Figure 19: The judgment for *The girls saw one movie each* when distributivity is not satisfied

The above judgment of a sentence containing binominal *each* is summarized in Table 5. The top header and the lefthand header represent the truth values of distributive quantification and the monotonicity constraint, respectively. The other cells between the two headers represent the judged truth values resulting from different combinations of the two components. $\star$ indicates undefinedness.

ψ \ ϕ	$\mathbb{T}$	$\mathbb{F}$
$\mathbb{T}$	$\mathbb{T}$	$\mathbb{F}$
$\mathbb{F}$	$\star$	$\mathbb{F}$

Table 5: The judged truth values of a sentence with binominal *each*

The discrepancy between the empirically attested truth value judgment and the theoretically predicted truth condition is not unique to the present account: all accounts that take the variation requirement to follow from some form of definedness condition suffer from this discrepancy. Note that treating the monotonicity condition as an at-issue component is not a tenable remedy: although it correctly predicts that failure to satisfy distributivity leads to falsity, it fails to predict the unacceptability induced by the failure of the monotonicity condition when distributivity is satisfied.

I do not have a solution to the discrepancy at this point. However, I would like to suggest some possible directions in which a solution may lie. To begin with, it is possible that an

---

[46]I thank a reviewer for helping to bring this issue to the forefront.

output context constraints is conditional in nature, in the sense that it is only checked when the output context is not an empty set. There is an intuitive appeal to this move: there is no point in checking an empty output context to ensure that it encodes a particular form of dependency. Second, we may even step back and ask if there indeed is a discrepancy in the first place. We are led to think that there is a discrepancy because I assimilated an output context constraint to a familiar kind of definedness condition, i.e., *a presupposition*. It is indeed true that for a presupposition, (assuming a simple picture) failure of satisfaction typically yields undefinedness. However, there is no reason that an output context constraint should behave in exactly the same manner as a presupposition (see also Kuhn (to appear) for an early taxonomy of output context constraints suggesting that there is considerable heterogeneity in how they behave). Perhaps it is normal for an output context constraint to be only operational when the output context is nontrivial. The exact nature of an output context constraint requires further investigation. However, regardless of what it turns out to be, it is likely that the monotonicity condition can be adapted to the new understanding.

6.2 Projection

An issue closely related to the status of the monotonicity constraint is its projection behavior. Interestingly, the projection behavior of the monotonicity constraint is quite different from that of presuppositions. In particular, it seems that the constraint never projects beyond any operator that scopes above distributivity. To see this, let us first consider a classical case of presupposition projection.

(76) a. If Paula saw Matt again, she would tell him the truth.
      b. Paula didn't see Matt again.
          Projected: Paula saw Matt at least once before.

As is well known, *again* is associated with a presupposition of prior occurrence and the presupposition projects beyond the antecedent of a conditional, as in (76-a), and beyond negation, as in (76-b) (Karttunen 1973). However, when binominal *each* is embedded in the antecedent of a conditional or inside the scope of negation, the monotonicity condition does not project. For example, there is no inference that the girls spotted different mistakes in (77-a) or (77-b).[47]

(77) a. If the girls spotted one mistake each, they should get a reward.
      b. The girls didn't spot one mistake each.
          <u>Not</u> projected: The girls spotted different mistakes.

The discrepancy between an input context constraint (i.e., a presupposition) and an output context constraint is not that surprising. After all, the former constrains *input* contexts while the latter constrains *output* contexts. If the monotonicity condition is a constraint on output contexts brought about by a distributive update, it is in fact predicted to only 'project' as far as the distributive update. In other words, it does not project beyond the scope of an operator embedding a distributive update.

---

[47](77-b) is in fact ambiguous between two readings. In the first reading, the numeral phrase *one mistake* takes scope over negation but falls in the distributive scope of the plural subject *the girls*. In this reading, the monotonic measurement condition is attested and requires that the girls did not spot the same mistake. In the second reading, the numeral phrase takes scope below negation. Unlike with the other reading, the monotonic measurement condition is no longer attested in this case. It is in this sense that the monotonic measurement condition does not project beyond negation.

7 Remarks on plurality measurement

For a sentence with binominal *each* to be well-formed, the monotonicity account makes the prediction that it must be possible to measure the relevant singular entities *and* the pluralities they form. However, measuring pluralities is not always straightforward. This section takes up less transparent cases involving plurality measurement, discusses the potential challenges they pose to the monotonicity account, and offer ways to handle the challenges.[48]

7.1 The need for mediating events

A measure function like weight measurement (μ_{weight}) involves a monotonic mapping from the individual domain to the measurement domain. Suppose that the measure function involved in (78) is μ_{duration}, i.e., duration measurement. It is not immediately clear that it is also such a monotonic mapping.

(78)   The boys worked for 10 hours each.

After all, μ_{duration} does not directly measure boys. Instead, it measures the duration of some working events with the boys as their agents. Hence, to account for the well-formedness of (78), it is then necessary to assume that (under a distributive reading) there is some kind of *working event* introduced for each boy and it is the events that are being measured for duration, as shown in the first two conjuncts in (79). The monotonicity condition then requires that the number of boys and the duration of their working events are monotonic—that is, more boys work for longer hours, as indicated by the third conjunct in (79).

(79) $\max^x(\text{boy } x) \wedge \delta_x(\exists e \wedge \text{work } e \wedge \text{Ag } e = x \wedge \mu_{\text{dur}} e = 10\text{hrs}) \wedge \text{dm}_{x,e} \, \mu_{\text{dur}}$

It has been argued that thematic functions that mediate individuals and events are homomorphic, i.e., structure-preserving with respect to the summation operation (Krifka 1992).[49] This means that in a distributive reading, there is a homomorphic mapping from a set of individuals and their pluralities to the events they participate in. As a result, the duration of the working events increases as the number of boys involved increases, satisfying monotonicity.[50]

7.2 'Proxy' measurement

Although duration measurement is monotonic, the closely related measurement of age does not seem to be monotonic. This can be seen by comparing the more well-formed (78) and the degraded (80).

(80)   ?The children are 10 years old each.

The status of (80) can be explained if age measurement is non-monotonic relative to individuals: Two 11-year-olds are not old enough to buy alcohol if the minimum age requirement for that is 21.[51]

---

[48]Many of the issues discussed in this section are credited to a reviewer of this paper.

[49]Given the standard assumption that the domains of individuals and events are partially ordered, thematic functions are also monotonic.

[50]For events with overlapping temporal intervals, collective duration can be measured along the lines of Krifka (1998).

[51]Note that age measurement is monotonic relative to time, as more time means greater age.

Although the age measurement in (80) is non-monotonic relative to individuals, it does not mean that the measure phrase '10 years old' has no chance of associating with any monotonic measure function. Previous studies have shown that measure phrases underdetermine measure functions (e.g., Schwarzschild 2005, Wellwood 2015). Notably, Schwarzschild (2005) shows that measure phrases have nonce uses and can be associated with non-canonical measure functions. For example, *five prisoners* usually measures the number of prisoners, but it can also measure their height, as in *the pile was five prisoners high* (cited from Schwarzschild 2005).

In light of this background, the measure phrase in (80) may be associated with a monotonic measurement if it can be construed as measuring another dimension that tracks the size (or number) of individuals. Since age often determines physical bulk, it is predicted that when the measure phrase *10 years old* goes proxy for bulk measurement, as indicated in (81), the naturalness improves.

(81)   The children are ten years old each. The full bed isn't big enough for them.

This is because bulk (measured in years lived) is homomorphic (and hence monotonic) to the number of individuals (before certain age).

Similarly, if people can be summed up based on their life experience (or wisdom) rather than their physical bulk and if life experience can be somehow standardized and measured in years, then it is predicted that one can more easily say:

(82)   These people are 50 years old each. They should know well enough that quarreling with each other won't get them anything.

This is because the measurement of life experience (measured in years lived) is again monotonic to the number of people.

7.3   Hidden dimension

Sometimes, the dimension can be tricky to determine. In the following example offered by a reviewer,

(83)   The cookies are 10cm each in diameter.

there seems to be no 'collective diameter' for the plurality consisting of several cookies. However, there is collective *length* for several diameters, which can be found by adding up the diameters of the cookies. If the involved measure function (or more accurately the dimension) is simply length, then additivity is not an issue here.

7.4   Loudness and brightness

Not all measurements are additive in a straightforward way. On the surface, we cannot simply add up the noise levels of two speakers: we can't say two speakers each of which is blasting 50db together gives out 100db. However, it is not because of the inability to add up sound pressures but because sound pressures are computed using logarithmic scales. The two speakers, when blasting sound waves not interacting with each other, are additive and give rise to an output of 53dB. More concretely, dB = $10\log(\frac{\text{power}}{\text{reference power}})$. Assuming power = 100000 and reference power = 1, we get 50dB. Now, if there is twice the power with two speakers with the same reference power, then we get 53dB (i.e., $10\log(200000)$). This shows us that in principle sound wave measurement is additive. For this reason, it is predicted that (84) is acceptable.

(84)   Those two speakers are blasting 50db each.

However, what about speakers that are placed far away so a person cannot hear both speakers at the same time? If a person moves close to a speaker and reports 50db of loudness from it, and moves away from it to get closer to another speaker, of which they also report 50db of loudness. If my understanding of the reviewer is correct, those who can accept (84) in the first scenario would still accept it in this scenario. If this is indeed the case, then we have to conclude that plurality measurement is not as simple as measuring two physical entities at a specific time and location. One may need to add an intensional component to plurality measurement to allow for measuring entities concatenated in a non-actual space and time.[52]

Somewhat relatedly, a reviewer showed that color shades are compatible binominal *each*:

(85)   This weekend I whitened my front teeth by two shades each.

The measure function involved in (85) is likely not perceived whiteness but perceived *change* in whiteness. If this is correct, then it is similar to the temperature change example discussed in (19-b) and (19-c). A change in color shade can perhaps be measured by how much whiteness (light reflected by an object) has been added to an object as compared to its previous color. So, if there are two (upper) front teeth being lightened and each of them is lighted by two shades. Then together the change in the light reflected by the two teeth is bigger than the change in the light reflected by just one of the two teeth.

7.5   Percentage measurement

Among the challenging examples raised by the reviewer, there is a specific subset concerning percentage:

(86)   a.   The Dow and the S&P gained 3 points each.
         b.   The governor has given his staff members raises of up to 20 percent each.

The challenges posed by these examples are two-fold. First, the percentages may be computed with different base numbers. For example, Employee A may have a salary of 100K and Employee B a salary of 200K. If A got a 10% raise and B a 15% raise. Their collective raise lies somewhere between 10% and 15% (i.e., 13.3%). It is predicted to not satisfy the monotonicity constraint. Second, even if the percentages are computed with the same base, say 100K, they seemingly fail to be monotonic. If both A and B got a raise of 10%, their collective raise is still 10%. This is because the percentage of their collective raise is computed based on their collective salary. Since they increase in the same proportion, the ratio does not change.

It's not difficult to see how this challenge can be resolved. The problem disappears if **absolute change** rather than relative change is considered. If a husband and a wife each get a 10% raise from their employer, they are able to bring more money home than if just one of them gets a 10% raise. So, if the proportional change in these examples can somehow be converted to absolute change, then binominal *each* is licensed. For concreteness, in (86-a) the monotonicity condition requires that the absolute change of the weighted average stock value in the two indices is bigger than the relevant absolute change in each index. In (86-b) it is required

---

[52]A potentially related example, also provided by the same reviewer, shows that one can talk about 'collective' impact (just like 'collective' noise):

(i)   The two outboard crewmen sustained 25 g's each at impact.

that the absolute monetary value associated with the two raises is bigger than each raise.[53] How exactly the suggestion here should be implemented is reserved for future research.

7.6 Remaining challenges from non-monotonic measurement

The last set of potentially problematic examples are the hardest to explain away. According to a reviewer, the following examples are all well-formed:

(87)   Speed
        a.   'All the fans work and are pushing 4000+/- rpms each.' (or: 'Not to mention scores of tiny plastic fans spinning at thousands of RPMs each.'
        b.   'This program says they're spinning at around 2000 rpm each.')

(88)   Air pressure
        The balloons were inflated to 12 atmospheres each.

(89)   Alcoholic density
        'Its Cosmopolitan and Margarita ice pops contain 6.41 percent alcohol each.'

(90)   Resolution
        '...the current displays for Amazon.com's Kindle Fire and Barnes & Noble's Nook, which sport tech specs of 169 pixels per inch each.'

If these examples are indeed well-formed, they pose a more serious threat to the monotonicity condition than previous examples considered earlier in this section. This is because it is not clear that any of the previous remedies readily apply to these examples, with the exception of proxy measurement.

    If we are to handle these examples as involving proxy measurement, we could say that the measurement involved in (87) can be reanalyzed as the amount of air being pushed (measured in rmp), or simply the number of rounds the fans can make together per minute, but crucially not the spinning speed. Both measurements are monotonic: more fans push more air; more fans make more rounds per minute (but not per fan). Similarly, what is being measured could be argued to be the amount of air in the balloons (and not the air pressure) in (88), the amount of ethanol (and not the concentration) in (89), and the screen size (and not screen resolution) in (90). All the proxy measurements are, again, monotonic.

    However, this line of reasoning runs the risk of rendering the monotonicity condition unfalsifiable, as also pointed out by a reviewer. If all non-monotonic measurements can be reanalyzed as proxy, monotonic measurements, then why do we observe ungrammaticality at the first place with non-monotonic measurements? To maintain the feasibility of proxy measurements, it is necessary to answer why proxy measurement is easier with some measurements but not others.

    One possibility is to tie the availability of proxy measurement to the commonality of a measure phrase: if a measure phrase is very common (like 20 km per hour, 50 miles per hour, or 70 degrees Fahrenheit, it might be strongly tied to a specific measure function. For this reason, it is harder to shift it to another measure function for proxy measurement to be possible. How far this correlation takes us will need to be clarified with more research.

---

[53]I suspect the same explanation holds for the range example:

(i)   The dynamic range and the temporal range were improved by four orders of magnitude each.

Alternatively, it may be useful to study closer the well-formedness of the examples in (87) – (90). It has been suggested to me that although (91) is similar to (90) in invoking resolution measurement, it is more degraded:

(91)   ?The displays of these tablets are 169 pixels per inch each.   (Simon Charlow, p.c.)

It is unclear why the same measurement fares better in (90) but worse in (91). To the extent that the contrast is a robust one, more factors need to be considered in the licensing of binominal *each*.

8 Conclusion

Borrowing an insight from previous studies that distributivity makes available functional dependencies with a nontrivial internal structure (Schein 1993, Lasersohn 1995, Krifka 1996b, van den Berg 1996, Landman (2000), Nouwen 2003, Brasoveanu 2008, Champollion 2010, 2017), this paper highlights the mereological nature of such dependencies. On the empirical front, it is shown that the mereological nature of distributive dependencies supports a monotonicity constraint, which subsumes three puzzling selectional requirements of binominal *each* reported in the literature (Safir and Stowell 1988; Sutton 1993; Zimmermann 2002; Zhang 2013). On the theoretical front, this paper proposes a novel version of dynamic plural logic that combines merits from van den Berg (1996)'s DPlL (i.e. dependencies are not introduced by default) and Brasoveanu (2008)'s PCDRT (i.e. domain pluralities are introduced in the range of assignment functions).

References

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy 9*(1), 5–16.

Balusu, R. (2005). Distributive reduplication in Telugu. In C. Davis, A. R. Deal, and Y. Zabbal (Eds.), *Proceedings of North East Linguistic Society 36*, Amherst, MA, pp. 39–53. University of Massachusetts GSLA Publications.

Barker, C. (1999). Individuation and quantification. *Linguistic Inquiry 30*(4), 683–691.

Barker, C. (2002). Continuations and the nature of quantification. *Natural Language Semantics 10*(3), 211–242.

Beck, S. (2010). Quantifiers in than-clauses. *Semantics & Pragmatics 3*, 1–72.

Beghelli, F. (1995). *The Phrase Structure of Quantifier Scope*. Ph. D. thesis, UCLA, Los Angeles.

Beghelli, F., D. Ben-Shalom, and A. Szabolcsi (1997). Variation, distributivity, and the illusion of branching. In *Ways of Scope Taking*, pp. 29–69. Kluwer Academic Press.

Beghelli, F. and T. Stowell (1997). Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi (Ed.), *Ways of Taking Scope*, pp. 71–107. Kluwer.

Bittner, M. (2014). *Temporality: Universals and Variation*. Oxford: Wiley Blackwell.

Brasoveanu, A. (2007). *Structured nominal reference and modal reference*. Ph. D. thesis, Rutgers University.

Brasoveanu, A. (2008). Donkey pluralities: Plural information states versus non-atomic individuals. *Linguistics and Philosophy 31*, 129–209.

Brasoveanu, A. (2011). Sentence-internal *different* as quantifier-internal anaphora. *Linguistics and Philosophy 34*, 93–168.

Brasoveanu, A. (2013). Modified numerals as post-suppositions. *Journal of Semantics 30*, 155–209.

Burzio, L. (1986). *Italian Syntax*. Reidel Publishers.

Cable, S. (2014). Distributive numerals and distance distributivity in Tlingit. *Language 90*, 562–606.

Champollion, L. (2010). *Parts of a whole: Distributivity as a bridge between aspect and measurement*. Ph. D. thesis, University of Pennsylvania, Philadelphia, Pennsylvania.

Champollion, L. (2015). Every boy bought two sausages each: Distributivity and dependent numerals. In U. Steindl, T. Borer, H. Fang, A. G. Pardo, P. Guekguezian, B. Hsu, C. O'Hara, and I. C. Ouyang (Eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, Somerville, MA, pp. 103–110. Cascadilla Proceedings Project.

Champollion, L. (2016). Covert distributivity in algebraic event semantics. *Semantics & Pragmatics 9*, 1–66.

Champollion, L. (2017). *Parts of a Whole: Distributivity as a Bridge between Aspect and Measurement*. Oxford: Oxford University Press.

Charlow, S. (2014). *On the semantics of exceptional scope*. Ph. D. thesis, New York University.

Charlow, S. (2020). The scope of alternatives: Indefiniteness and islands. *Linguistics and Philosophy 43*, 427–472.

Charlow, S. (to appear). Post-suppositions and semantic theory. *Journal of Semantics*.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry 37*(4), 535–590.

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press.

Chierchia, G., D. Fox, and B. Spector (2011). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. Maienborn, K. von Heusinger, and P. a. Portner (Eds.), *Handbook of Semantics*, Volume 3.

Choe, J.-W. (1987). *Anti-quantifiers and a theory of distributivity*. Phd, University of Massachusetts-Amherst.

Cresti, D. (1995). Extraction and reconstruction. *Natural Language Semantics 3*(1), 79–122.

de Swart, H. E. (2000). Scope ambiguities with negative quantifiers. In K. von Heusinger and U. Egli (Eds.), *Reference and anaphoric relations*, pp. 109–132.

DeVries, K. (2016). *Independence Friendly Dynamic Semantics: Integrting exceptional scope, Anaphora and their interactions*. Ph. D. thesis, University of California at Santa Cruz.

Dotlačil, J. (2010). *Anaphora and distributivity: a study of same, different, reciprocals and others*. Phd. dissertation, University of Utrecht.

Dotlačil, J. (2012). Binominal *each* as an anaphoric determiner: a compositional analysis. In *Proceedings Sinn und Bedeutung 16*.

Dotlačil, J. and R. Nouwen (2016). The comparative and degree pluralities. *Natural Language Semantics 24*, 45–78.

Farkas, D. (1997). Dependent indefinites. In F. Corblin, D. Godard, and J.-M. Marandin (Eds.), *Empirical issues in formal syntax and semantics*, pp. 243–268. Bern: Peter Lang Publishers.

Farkas, D. (2002a). Specificity distinctions. *Journal of Semantics 19*, 213–243.

Farkas, D. F. (2002b). Varieties of indefinites. In B. Jackson (Ed.), *Proceedings of SALT XII*, Cornell University, Ithaca, NY, pp. 59–83. CLC Publications.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland and P. Statava (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. Plagrave Macmillan.

Gil, D. (1988). Georgian reduplication and the domain of distributivity. *Linguistics*, 1039–1065.

Groenendijk, J. and M. Stokhof (1991). Dynamic predicate logic. *Linguistics and Philosophy 14*, 38–100.

Hackl, M. (2000). *Comparative Quantifiers*. Ph. D. thesis, Massachusetts Institute of Technology.

Heim, I. (1983). On the projection problem for presuppositions. In D. P. Flickinger (Ed.), *Proceedings of WCCFL 2*, Stanford University, Stanford, California, pp. 114–125. CSLI Publications.

Heim, I. and A. Kratzer (1998). *Semantics in generative grammar*. Malden, MA: Blackwell.

Henderson, R. (2014). Dependent indefinites and their post-suppositions. *Semantics & Pragmatics 7*, 1–58.

Huang, S.-Z. (1996). *Quantification and Predication in Mandarin Chinese*. Phd. dissertation, University of Pennsylvania.

Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in english. *Natural Language and Linguistic Theory 14*(2), 305–354.

Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk (Ed.), *Formal Methods in the Study of Language*. Amsterdam: Mathematical Center.

Karttunen, L. (1973). Presupposition of compound sentences. *Linguistic Inquiry 4*(3), 169–193.

Kayne, R. (2015, June). English *one* and *ones* as complex determiners. New York University.

Kennedy, C. (2015). A "de-Fregean" semantics (and neo-gricean pragmatics) for modified and unmodified numerals. *Semantics & Pragmaticss 8*, 1–44.

Kennedy, C. and J. Stanley (2009). On *average*. *Mind 118*, 583–646.

Kratzer, A. (2007). On the plurality of verbs. In J. Dölling, T. Heyde-Zybatow, and M. Shafer (Eds.), *Event structures in linguistic form and interpretation*, pp. 269–300. Berlin: Mouton de Gruyter.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsh, J. Van Benthem, and P. van Emde Boas (Eds.), *Semantics and Contextual Expression*, Dordrecht, Netherlands, pp. 75–115. Foris.

Krifka, M. (1991). A compositional semantics for multiple focus constructions. In S. Moore and A. Z. Wyner (Eds.), *Proceedings of the First Semantics and Linguistic Theory Conference*, Cornell University, pp. 127–158. Cornell University Working Papers in Linguistics.

Krifka, M. (1992). Thematic relations as links between nominal reference. In I. A. Sag and A. Szabolcsi (Eds.), *Lexical Matters*, pp. 29–54. Stanford, California: Center for the Study of Language and Information.

Krifka, M. (1995). Common nouns: a contrastive analysis of English and Chinese. In *The Generic Book*.

Krifka, M. (1996a). Parametrized sum individuals for plural anaphora. *Linguistics and Philosophy 19*(6), 555–598.

Krifka, M. (1996b). Pragmatic strengthening in plural predications and Donkey sentences. In T. Galloway and J. Spence (Eds.), *Proceedings of SALT VI*, Ithaca, NY, pp. 136–153. CLC Publications.

Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and Grammar*, pp. 197–236. Dordrecht: Kluwer Academic Publishers.

Kuhn, J. (2015). *Crosscategorial singular and plural references in sign language*. Ph. D. thesis, New York University.

Kuhn, J. (2017). Dependent indefinites: the view from sign language. *Journal of Semantics 34*, 407–446.

Kuhn, J. (to appear). The dynamics of negative concord. *Linguistics and Philosophy*.

Landman, F. (2000). *Events and plurality: The Jerusalen Lectures*, Volume 76 of *Studies in Linguistics and Philosophy*. Dordrecht: Kluwer.

Lasersohn, P. (1995). *Plurality, Conjunction and Events*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lauer, S. (2009). Free relatives with *-ever*: Meaning and use. Ms. Stanford University.

Lauer, S. (2012). Some news on *Irgendein* and *algún*. Ms. Stanford University.

Law, J. H.-K. (2019). *Constraints on distributivity*. Ph. D. thesis, Rutgers University, New Brunswick, NJ.

Link (1987). Generalized quantifiers and plurals. In P. Gärdenfors (Ed.), *Generalized Quantifiers. Linguistic and Logical Approaches*, Volume 31 of *Studies in Linguistics and Philosophy*, pp. 151–180. D. Reidel Publishing Company.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, and A. v. Stechow (Eds.), *Meaning, Use and Interpretation of Language*. Berlin: DeGruyter.

Lønning, J. T. (1987). Mass terms and quanti cation. *Linguistics and Philosophy 10*(1), 1–25.

May, R. (1985). *Logical Form: Its Structure and Derivation*. Cambridge, Massachusetts: MIT Press.

Milačić, D., R. Singh, and I. Toivonen (2015). On the morphosyntactic representation of dependent quantification: Distance distributivity, dependent indefinites, and Skolemization. In E. Csipak and H. Zeijlstra (Eds.), *Proceedings Sinn und Bedeutung 19*, Göttingen, Germany., pp. 411–425.

Moltmann, F. (1991). On the syntax and semantics of binary distributive quantifiers. In *Proceedings of the North East Linguistic Society (NELS) 22*, pp. 279–292.

Muskens, R. (1996). Combining Montague semantics and discourse representation. *Linguistics and Philosophy 19*(2), 143–186.

Nouwen, R. (2003). *Plural pronominal anaphora in context: Dynamic aspects of quantification*. Ph. D. thesis, Utrecht Univeristy, Utrecht, The Netherlands.

Perlmutter, D. M. (1970). On the article in English. In M. Bierwisch and K. E. Heidolph (Eds.), *Progress in Linguistics*, pp. 233–248. The Hague: Mouton.

Postal, P. M. (1974). On certain ambiguities. *Linguistic Inquiry 5*(3), 367–424.

Rett, J. (2014). The polysemy of measurement. *Lingua 143*, 242–266.

Roberts, C. (1987). *Modal Subordination, Anaphora and Distributivity*. Ph. D. thesis, University of Massachusetts, Amherst.

Safir, K. and T. Stowell (1988). Binominal *each*. In J. Blevins and J. Carter (Eds.), *Proceedings of North East Linguistic Society 8*, Toronto, pp. 426–450. GLSA.

Schein, B. (1993). *Plurals and events*. Cambridge, Massachusetts: MIT Press.

Schwarzschild, R. (1996). *Pluralities*. Dordrecht ; Boston: Kluwer Academic.

Schwarzschild, R. (2002). The grammar of measurement. In B. Jackson (Ed.), *Proceedings of SALT XII*, Cornell University, Ithaca, NY, pp. 225–245. CLC Publications.

Schwarzschild, R. (2005). Measure phrases as modifiers of adjectives. *Recherches linguistiques de Vincennes 34*, 207–228.

Schwarzschild, R. (2006). The role of dimensions in the syntax of noun phrases. *Syntax 9*(1), 67–110.

Scontras, G. and N. D. Goodman (2017). Resolving uncertainty in plural predication. *Cognition 168*, 294–311.

Solomon, M. (2011). True distributivity and the functional interpretation of indefinites. Unpublished manuscript.

Solt, S. (2015). Q-adjectives and the semantics of quantity. *Journal of Semantics 32*, 221–273.

Solt, S. (2016). On measurement and quantification: The case of *most* and *more than half*. *Language 92*, 65–100.

Stowell, T. (2013). Binominal *Each*: A dp that may not be. In D. Gil, S. Harlow, and G. Tsoulas (Eds.), *Strategies of quantification*, Oxford, pp. 260–313. Oxford University Press.

Strawson, P. F. (1950). On referring. *Mind 59*, 320–344.

Sutton, M. (1993). Binominal *each*. Master's thesis, UCLA.

Szabolcsi, A. (2010). *Quantification*. Cambridge University Press.

van den Berg, M. (1996). *Some aspects of the internal structure of discourse*. Ph. D. thesis, University of Amsterdam, Amsterdam.

Wellwood, A. (2015). On the semantics of comparison across categories. *Linguistics and Philosophy 38*, 67–101.

Zhang, L. (2013). A ratio analysis of binominal each. In *The 3rd Mid-Atlantic Colloquium of Studies in Meaning (MACSIM 3)*.

Zimmermann, M. (2002). *Boys buying two sausages each: On the syntax and semantics of distance-distributivity*. Phd, University of Amsterdam.