CONSTRAINTS ON DISTRIBUTIVITY

By

HOI KI LAW

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Linguistics

Written under the direction of

Simon Charlow and Veneeta Dayal

And approved by

____

____

____

____

New Brunswick, New Jersey

October, 2019

ABSTRACT OF THE DISSERTATION

Constraints on distributivity

By HOI KI LAW

Dissertation Director:

Simon Charlow and Veneeta Dayal

Distributivity can be marked with lexical items like binominal *each* in English:

(1) The girls read three books each.

It has long been noted that some distributivity markers need to be licensed by the morphosyntactic makeup and/or interpretive properties of the predicate being distributed over (e.g, read three books in (1)). In this dissertation I investigate three distributivity markers that exhibit this type of licensing requirement:

1. binominal *each* in English (Safir and Stowell 1988, Zimmermann 2002, Champollion 2015, Kuhn 2017);
2. the verbal distributivity suffix *saai* in Cantonese (Tang 1996, Lee 2012);
3. the adverbial distributivity marker *ge* in Mandarin (Kung 1993, Lin 1998b, Lee et al. 2009a).

The investigation leads to two findings. The first finding is that these licensing requirements should be understood as constraints on the dependencies arising from distributive quantification, which echo similar constraints proposed for various types of indefinites (Farkas 1997, 2002b, Brasoveanu and Farkas 2011, Henderson 2014, Kuhn 2017). A consequence of this finding is a more general

conception of constraints on dependencies: they are not only associated with indefinites (as conceived of in Farkas (2002b)), as they may be borne by distributivity markers.

The second finding is that constraints on dependencies may differ along a few parameters. One parameter determines whether a constraint makes reference to the internal mereological structure of dependencies, which arise from evaluating distributivity. Using the interactions of distributivity markers with extensive and intensive measure phrases (Zhang 2013), I conclude that the constraints under investigation make reference to the mereological nature of distributive dependencies. These constraints stand in contrast with constraints previously formulated for dependent indefinites (Farkas 1997, 2002b, Henderson 2014, Champollion 2015, Kuhn 2017), which do not need to access the mereological structure of dependencies. Another parameter determines whether a constraint requires dependence or independence. Using the contrast between binominal each and Mandarin ge on the one hand, and Cantonese saai on the other hand, I show that both parameters are used in natural language. This conclusion adds further support to the parallelism between constraints contributed by distributivity markers and those contributed by indefinites, as the dependence-independence parameter has also been used to characterize dependent indefinites and specific indefinites (e.g., Farkas 2002b, von Heusinger 2002).

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 constraints to act as output constraints on distributive quantification, which mirror the use of output constraints in studies like Farkas (1997, 2002b), Henderson 2014, and Kuhn (2017).

ACKNOWLEDGEMENTS

This dissertation can be roughly decomposed into dynamic semantics, distributivity, plurality, mereology, measurement, higher order meaning, cross-linguistic research, persistence (in a non-technical sense) and passion (in a non-technical sense). The people thanked below have helped me gain insight into one or more of these aspects.

Simon Charlow (co-chair of my dissertation committee) joined Rutgers in 2014 and started teaching beautiful courses on dynamic semantics and monadic compositional semantics. The word 'inspirational' is an understatement of the effect of his teaching. My husband Haoze Li and I were in the same courses and we often spent hours admiring the way he packaged a lesson to make complex ideas flow naturally from simple ones. Thanks to him, I have learned that complexity is a set of simplicities in disguise, and there is nothing more satisfying than being able to use semantic tools as lego blocks. The dynamic semantics and higher order meaning part of the dissertation are credited to his teaching and research.

Veneeta Dayal (co-chair of my dissertation committee) taught me kinds, plurals, quantifiers, and questions, all of which (except for questions) play an important part in this dissertation. In addition, I have learned not just to appreciate the value of cross-linguistic research, but also how to put together threads of phenomena to form a holistic, theoretically interesting picture. Had it not been for Veneeta, this dissertation would have been a lot dryer for the writer and the reader. I also would like to thank Veneeta for reminding me to be open-minded and not be a fan of any framework. I have found this advice extremely liberating!

Ken Safir and Robert Henderson are both wonderful committee members to have. They have worked on topics that are directly related to the present work. Many of the key observations in the binominal *each* chapter go back to Ken's joint work with Tim Stowell. The semantic architecture in the whole dissertation is inspired by Robert's work on dependent indefinites. I'm grateful to them for the beautiful research they did and the comments and suggestions they have given to me throughout the writing of this dissertation. A fun fact: Ken is a semanticist in disguise, and Robert has even more guises, as a morphologist, syntactician, and phonologist. A learning point: maybe don't be a fan of any subfield.

I started working on Mandarin *ge* (the last chapter of this dissertation) after taking an extremely well-taught seminar on distributivity and event semantics with Lucas Champollion. His seminar touched upon not just distributivity, but also plurality and measurement, which are key players in this dissertation. Although I have not pursued an event-based framework for analyzing distributivity, there was a time when the core ideas of this dissertation lived in plural events rather than plural info-states. If not for the seminar with Lucas, I would not have gotten interested in distributivity at all, as I used to think that this topic was well explored enough.

I met with Anna Szabolcsi a few times to discuss research in this dissertation and those meetings have proven to be very important. An earlier version of the monotonicity constraint was too strong, as persistently pointed out by Anna (and less persistently by others). The weaker version proposed in the dissertation was developed following a meeting with Anna.

Kristen Syrett has to be praised for not shouting to me 'hello, I work on distributivity!' I only found out quite late in the game that I could have recruited more of her help. That said, through working with Kristen on other projects I have discovered that my heart is big enough for both theoretical linguistics and experimental linguistics. I cannot thank her enough for demonstrating to me how to fruitfully combine these two pursuits. Kristen's time management and persistence are inspirational. Every time I find it too hard to balance my role as a researcher and a mother, I think of Kristen, her children, her dogs, her yoga, her teaching, her research, and her passions. Peace always comes to me afterwards.

I'm also indebted to Yimei Xiang for her friendship and support, to Jane Grimshaw for her inspiring teaching, to Maria Bittner for her courses in dynamic semantics, as well as to Akin Akinlabi and Paul de Lacy for the phonology courses I have taken with them.

Many fellow graduate students at Rutgers Linguistics have helped me adapt to the life in New Jersey. Sarah Hansen, Diti Bhadra, and Yağmur Sağ has provided me with the much needed friendship and company. Ümit Atlamaz and his lovely family always inspire me to become a better parent. Events with Kunio Kinjo, Shu-hao Shih, Mingming Liu, Sha Zhu and Haoze Li formed some of my fondest memories about Rutgers. Above all, I admire my friends for their passions for linguistic research. There is no problem too hard for them. I'm blessed to have been in this linguistic community.

The rich experience outlined above is only possible because the linguistics department at Chinese University of Hong Kong had cultivated my interest in linguistics and supported my early research in syntax (with Candice Cheung) prior to my study at Rutgers. I thank the courageous group (Gladys Tang, Virginia Yip, Thomas Lee, Yang Gu, Ping Jiang) for their efforts in establishing the first linguistics department in Hong Kong, which directly benefited students like me.

Lastly, I'm very grateful to my family for their constant support and patience. Without their unconditional support, I would not have been able to pursue a research career, let alone writing a dissertation on theoretical linguistics.

DEDICATION

To my parents: Siu King Liu and Ming Tung Law.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Overview

Research on distributivity has established that there are morphological markers across languages that signal distributive quantification (Roberts 1987, Link 1987, among many others). A well-known example is English *each*, as shown in (1).[1]

(1) The girls *each* made a kite.

Many generalizations about distributivity markers have been drawn that shed light on the process of distributive quantification. For instance, a requirement is that the subject being targeted by *each* for distributive quantification (i.e., the *distributivity key*) has to contribute a plural entity. This requirement distinguishes *the girls* in (1) from *the girl* in (2), as only the former contributes a plurality for quantification. This requirement informs us that distributive quantification must not be vacuous—should there be only one girl, then ordinary predication without distributivity already suffices.

(2) *The girl each made a kite.

[1] For the purpose of this study, 'signaling' distributivity does not equal contributing a distributivity operator. While some distributivity markers may indeed contribute a distributivity operator, other may merely require the presence of a distributivity operator.

As another example, some distributivity markers, like *each*, cannot quantify over a mass noun phrase, as shown in (3), suggesting that these distributivity markers only have the ability to break down pluralities into their atomic parts (Roberts 1987, Link 1987, Zimmermann 2002, Champollion 2010).[2]

(3) *The water *each* leaked to the floor.

There are many other observations of this kind, which give rise to rich generalizations ranging from cover distributivity (when distributive quantification is over a cover of the distributivity key; see Scha 1981, Schwarzschild 1996) to long-distance distributivity (when the distributivity marker is not adjacent to the distributivity key; see Zimmermann 2002, Dotlačil 2012, Champollion 2017). However, it is fair to say that the majority of these generalizations pertain to the relationship between a distributivity marker and a distributivity key.

Not only until more recently have linguists started to pay more attention to the relationship between distributivity markers and expressions in their *distributed share*, i.e., the predicate being distributively predicated of the distributivity key. A survey of the literature reveals that many distributivity markers impose selectional requirements on what expressions may show up in their distributed share and the range of interpretations these expressions may assume. However, many of the selectional requirements on the distributed share are still poorly understood.

The aim of this dissertation is to gain a better understanding of share requirements in general by investigating a few distributivity markers that have been reported to impose share requirements. The overarching finding, which I report in the next few chapters, is that share requirements are **constraints on functional dependencies**.[3] More precisely, share requirements are constraints on the functional dependencies a distributivity share participates in. This conclusion provides corroborations for two lines of research in the linguistic literature.

[2]There are many distributivity markers that do not require a distributivity key with accessible atomic parts. A famous example is Mandarin *dou* (Lin 1998a). However, *dou*'s status as a distributivity marker is disputed (Xiang 2008, Liu 2016, Xiang 2016). English *all* is another often cited example, but its status has also been challenged by Brisson (1998, 2003) (cf. Champollion 2017).

[3]I use the term 'functional dependency' to refer to two sets X and Y that stand in a certain relationship R. If R does not map the same individual in X to different values in Y (in other words, if R is a function), then Y we can naturally say that Y is functionally dependent on X. If R is a relation that maps a single value in X to more than one values in Y, then a functional dependency can still be obtained by defining $f : X \to \mathcal{P}(Y)$. For concreteness, suppose $X = \{\mathsf{a}, \mathsf{b}\}$, $Y = \{\mathsf{c}, \mathsf{d}, \mathsf{e}\}$ and they stand in a relation $R = \{\langle \mathsf{a}, \mathsf{c}\rangle, \langle \mathsf{a}, \mathsf{d}\rangle, \langle \mathsf{b}, \mathsf{e}\rangle\}$. We can define an $f : X \to \mathcal{P}(Y)$, such that $f\mathsf{a} = \{\mathsf{c}, \mathsf{d}\}$ and $f\mathsf{b} = \{\mathsf{e}\}$. This general definition of functional dependencies can be used for a functional structure or a relational structure.

First, it provides additional evidence that distributive quantification brings about a set of functional dependencies and these dependencies should be made accessible to compositional semantics. Roberts (1987), Schein (1993), Kamp and Reyle (1993), Lasersohn (1995), Elworthy (1995), Krifka (1996a), van den Berg (1996), and Jackendoff (1996) are some of the studies that argue for the representation of the functional dependencies arising from distributivity. Since then, various attempts have been made to make available these dependencies, mainly using resources from Events Semantics (Schein 1993, Lasersohn 1995, Champollion 2017) and various versions of Dynamic Semantics (Elworthy 1995, Krifka 1996a, van den Berg 1996, Nouwen (2003), Brasoveanu 2008).[4]

Second, it narrows the gap between garden-variety distributivity markers and markers of distributive numerals. Distributive numeral markers have long been noted to place a dependence requirement on distributivity, as argued in Choe (1987a), Farkas (1997), Balusu (2005), Henderson (2014), Cable (2007), and Kuhn (2017). If the findings from this dissertation are to stay, there is no fundamental difference between many distributivity markers and markers of distributive numerals: they both signal distributive quantification and impose constraints on its outcome. Building on the dependence requirement of distributive numerals, this dissertation takes up a variety of novel requirements distributivity markers in various languages impose on the functional dependencies of distributivity.

The distributivity markers taken up in this research are English *each* (in three positions, binominal in (4-a), adverbial in (4-b), and determiner in (4-c)), Cantonese *saai* (5), and Mandarin *ge* (6).

(4) a. The girls made one kite *each*.
    b. The girls *each* made a kite.
    c. Each girl made a kite.

(5) Di-neoizai zing-*saai* fungzeng.
    CL-girl make-SAAI kite
    'The girls each made one or more kites.' *Cantonese*

[4] Krifka (1996a), van den Berg (1996), Nouwen (2003), and Brasoveanu (2008) are extensions of DPL, while Elworthy (1995) is couched in Elworthy's own Theory of Anaphoric Information (TAI). For a useful comparison among Elworthy (1995), van den Berg (1996), Krifka 1996a, see Nouwen (2003). For a useful comparison between van den Berg (1996) and Brasoveanu (2008), see Brasoveanu (2007) and Chapter 2 of this dissertation.

(6) Nühai-men *ge* zuo-le yi-zhi fungzeng.
girl-PL GE make-ASP one-CL kite
'The girls each made a kite.' *Mandarin*

What makes these distributivity markers special is that they have all been remarked to impose additional morphosyntactic or interpretive requirements on their distributed share.

- Binominal *each* requires the support of so-called counting quantifiers in the distributed share (Safir and Stowell 1988, Sutton 1993, Szabolcsi 2010), and requires that the counting quantifier obligatorily co-vary with the distributivity key (Choe 1987a). *Each* in other positions have been reported to signal that the events participating in distributive predication are *disjoint events*, or differentiated events, in the terms of Tunstall (1998) (see also Beghelli and Stowell (1997), Brasoveanu and Dotlačil 2015).
- Cantonese *saai* generally resists counting quantifiers in the distributed share (Lee 1994, Tang 1996), unless they fail to co-vary with the distributivity key.
- Mandarin *ge* needs to be supported by a counting quantifier or an expression with a pronoun bound by the distributivity key (Kung 1993, Lin 1998b, 2005, Soh 2005, Tsai 2009, Lee et al. 2009a).

Despite having been documented and studied on a case-by-case basis, the share requirements of these distributivity markers have not been studied together as a **natural class of phenomena** pertaining to distributivity. It is useful to take them up as a natural class for two reasons. For one thing, we will get a more holistic picture of share requirements, which will help us understand their role in natural language. For another, since they are all found on markers that signal distributivity, we should expect them to be intimately related to distributive quantification. The following heuristic suggested in Szabolcsi (1997) is useful here.

> What range of expressions actually participates in a given process is suggestive of exactly what that process consists in.

Specifically, the range of share requirements that may show up with distributive quantification is suggestive of what distributive quantification consists in.

Despite the apparent heterogeneity of the share requirements studied in this dissertation, I argue that the share requirements have a common core—they are all constraints on the functional dependencies arising from distributive quantification. The differences lie in the type of constraints, which we can understand in terms of a few parameters.

It is useful to discuss, at an informal level, how functional dependencies can be constructed with help of a sentence like (7). Intuitively, this sentence can be verified in a number of ways. Two examples are shown in Figure 1.1.

(7) The girls (each) read a book.

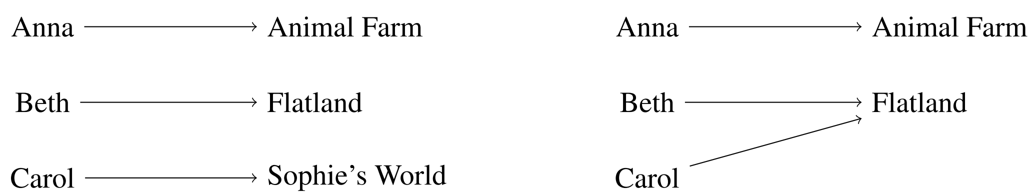


Figure 1.1: Some functional dependencies following the evaluation of (7)

In each scenario, a girl stands in the reading relation with a book. In the one on the left, every girl read a different book. In the one on the right, two girls read the same book. However, both scenarios verify (7). Let us call these informal representations of how the girls stand in relation to the books they read *functional dependencies*. A functional dependency is thus called because a functions can be defined with the domain set to one part of the relation (typically the distributivity key) and the range set to another part of the relation (typically information contributed by expressions in the distributed share).[5] A sentence with a distributivity marker may give rise to a set of functional dependencies, expressing relationships between the distributivity key and various expressions in the distributed share. With this much background, we are ready to discuss different types of constraints that are imposable on these functional dependencies.

First, they may differ on whether the functional dependency is required to exhibit *dependence*, as shown in Figure 1.2 (left) or *independence*, as shown in Figure 1.2 (right). The former requires that part of the distributed share co-varies with the distributivity key, as has been argued for distributive numerals and distributivity with binominal *each* in English, and will be argued for other uses of *each*

[5]See also footnote 3.

as well as distributivity with Mandarin *ge*. The latter leads to the lack of co-variation of part of the distributed share relative to the distributivity key, as will be argued for distributivity with Cantonese *saai*.

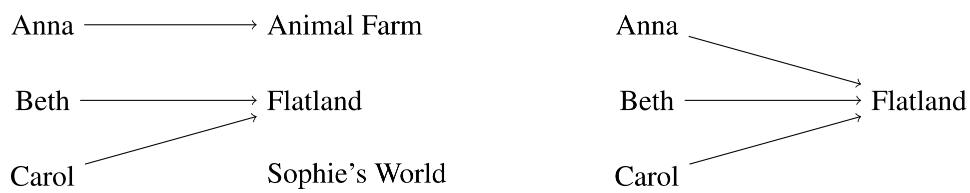


Figure 1.2: Functional dependencies that exhibit dependence (left) and independence (right)

Second, they may differ on whether (in)dependence is required at the *value* level or at the *structural* level. Informally speaking, a functional dependency is said to exhibit **value (in)dependence** when (in)dependence is determined without making reference to the internal mereological structure of the dependency. Otherwise, it exhibits **structural (in)dependence**.

A concrete example will help illustrate the difference. Suppose the interpretation of (8) establishes a functional dependency between a set of angles and their corresponding angle degrees, as shown in Figure 1.3 (left) and the interpretation of (9) does so for a set of drinks and their corresponding temperatures, as shown in Figure 1.3 (right). The two types of functional dependencies look exactly the same and lack value dependence.

(8) The angles are 60 degrees each.

(9) *The drinks are 60 degrees (Fahrenheit) each.

Figure 1.3: Both functional dependencies lack value dependence

However, the functional dependency on the left differs crucially from the one on the right in an important way. The former is built based on the an *extensive* measure function, which is *additive*, while the latter is built based on an *intensive* measure function, which is not additive. The definition of additivity given in (10) is taken from Krifka (1998) (with the concatenation operator there replaced by a summation operator here).

(10) A measure function μ (from a set of entities in D to a set of positive real numbers) is additive iff
$\forall x, y \in D : \mu(x \oplus y) = \mu(x) + \mu(y)$

The functional dependency built with an extensive measure function licenses inferences about the measurement of the mereological sum of all the individuals in the functional dependency. In addition, there is a guarantee that bigger individuals are mapped to bigger measurements (e.g., in terms of numerical values), as shown in Figure 1.4 (left). However, the functional dependency built with an intensive measure function cannot license the same type of inference. We need much more information to find out about the measurement of the sums, and there is no guarantee that bigger individuals will always be mapped to bigger degrees, as shown in 1.4 (right).

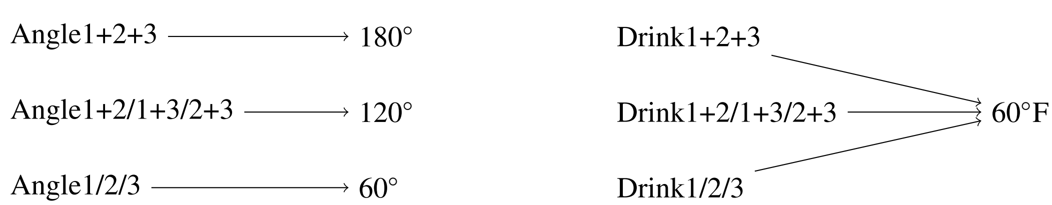


Figure 1.4: Functional dependencies with structural dependence (left) and without (left)

Constraints requiring value dependence have been explored in previous studies on distributive numerals (as well as dependent indefinites). In this study, I show that structural dependence is also relevant for formulating constraints on distributivity. In particular, I argue that English *each* and Mandarin *ge* exhibit structural dependence, while Cantonese *saai* exhibit structural independence. These claims are summarized in Table 1.1

In summary, this dissertation advocates to treat garden-variety distributivity markers along the lines of distributive numeral markers in making a two-part contribution: signaling distributivity and imposing constraints on the functional dependencies arising from distributive quantification. By doing so, it demonstrates that in addition to value dependence, a relatively well-studied constraint, there is a wider range of constraints that target functional dependencies contributed by distributivity.

	Value	Structure
Dependence	Distributive numerals	Binominal *each*
	Adverbial and determiner *each*	Mandarin *ge*
Independence	–	Cantonese *saai*

Table 1.1: Major empirical claims in the dissertation

Since having access to the functional dependencies of distributivity is of critical importance for modeling constraints on distributivity, I discuss how to build these functional dependencies in the next section.

1.2 Building functional dependencies

A functional dependency is a relation between two sets, such that determining a value in the first set uniquely determines a value in the second set. Many studies have observed that sentences with a distributive quantifier (*every NP* or *each NP*) or a distributivity marker give rise to functional dependencies. For example, consider the following data:

(11) Every man^x loves a woman^y. The old men$_x$ bring them$_y$ flowers to prove this. (van den Berg 1996:126)

(12) The students^x each wrote an article^y. They$_x$ each sent it$_y$ to L&P. (Krifka 1996a:557)

The plural pronoun in (11) has to refer to the *corresponding* women loved by the old men, as introduced in the preceding sentence. Likewise, the singular pronoun in (12) picks out a different article written by each student, a dependency introduced in the preceding sentence.

A few studies have motivated to extend DRT or other versions of dynamic semantics to model anaphora to dependency (e.g., Elworthy 1995, Krifka 1996a, van den Berg 1996, Nouwen 2003, Brasoveanu 2007, 2008). This study follows van den Berg (1996), Brasoveanu (2007, 2008), Henderson (2014) and uses Dynamic Plural Logic (DPlL) to model the functional dependencies established as a by-product of evaluating distributive quantification. I have chosen this framework for the following reasons:

- It is a plural logic capable of modeling and compositionally constructing functional dependencies, which are important for stating constraints on these dependencies.
- It is a dynamic logic capable of transmitting functional dependencies in the course of interpretation. These functional dependencies can be retrieved with use of standard anaphoric devices in dynamic semantics, allowing a streamlined compositional analysis.
- Other recent studies have used a version of DPlL for studying distributive numerals (e.g., Henderson 2014, Champollion 2015, Kuhn 2017). Having the present study couched in a similar framework enables us to compare the present work with these studies.

I introduce the framework of DPlL used in this dissertation in Chapter 2. The backbone of the logic comes from van den Berg (1996). On top of that, I borrow domain pluralities and compositionality from Brasoveanu (2007, 2008) (but not dependency-introducing variable introduction (aka. random assignment) and distributively-evaluated lexical relations). Since functional dependencies can be modeled with other frameworks, I discuss, in the second part of Chapter 2, some alternatives to using DPlL and why I have not adopted them in this dissertation.

1.3 Distributivity markers studied in this dissertation

1.3.1 Binominal *each*

Binominal *each* is taken up in Chapter 3. The key observation is that noun phrases marked by binominal *each* pattern like dependent indefinites in requiring obligatory co-variation with the distributivity key, as shown in (13) (Safir and Stowell 1988, Champollion 2015, Kuhn 2017). In addition, these noun phrases must be either counting quantifiers (Sutton 1993, Szabolcsi 2010) or measure phrases with an extensive measure function (Zhang 2013), as shown in (14) – (15-b).

(13) The girls read two books each, namely, *Brave New World* and *Animal Farm*.

(14) The girls read *some/two books each. — Counting quantifier

(15) a. The angles are 60 degrees each. — Extensive measurement

b. *The drinks are 60 degrees each. — Intensive measurement

These properties are used to motivate a constraint requiring structural dependence in the functional dependencies established by distributivity with binominal *each*. The constraint is called a **monotonic measurement constraint** to indicate the importance of measurement in the formulation of this constraint.

1.3.2 Cantonese *saai*

Saai is a verbal suffixal serving to mark distributivity in Cantonese (Lee 1994, Tang 1996, Lee 2012). *Saai* has been remarked to resist indefinites in the distributed share, as shown in (16), but not definites, as shown in (17). Although previous studies have speculated that *saai* has a preference for specificity or definiteness (Lee 1994, Tang 1996), it has not been made clear how such a preference is related to *saai*'s role as a distributivity marker. In addition, *saai* also displays resistance to extensive measurement. Using a measure phrase ambiguous between an extensive measurement reading and an intensive measurement reading, such as *saam-sing-ge seoi* 'three liters of water, three-liter water' in (18), it can be shown that *saai* is only compatible with the latter (see also Schwarzschild 2006).[6]

(16) *Di-neoizai zing-*saai* jat-zek fungzeng.
CL.PL-girl make-SAAI one-CL kite
Intended: 'The girls each made one or more kites.' Indefinite

(17) Di-neoizai gin-*saai* go-go lousi.
CL.PL-girl make-SAAI one-CL kite
'The girls each saw the teacher.' Definite

(18) Di-neoizai maai-saai saam-sing-ge seoi.
CL.PL-girl buy-SAAI three-liter-MOD water
a. *The girls each bought three liters of water. Extensive
b. The girls each bought three-liter water. Intensive

In Chapter 4, I show that there is a principled reason for the resistance to indefinites–*saai* requires the functional dependency between the distributivity key and the part of the share contributed by a post-*saai* expression to exhibit *independence*. Sentences like (16) are ruled out because when a kite is chosen independently of the boys, there could only be one kite, and the boys could not

[6] *Saai* is incompatible with most measure predicates so minimal pairs like (15-a) and (15-b) cannot be constructed.

distributively make the same kite. I explore two analyses to account for *saai*'s resistance to dependence, and ultimately argue in favor of an **independence constraint**. The constraint is shown to stand in opposition to the structural dependence requirements (i.e., the monotonicity constraint) of binominal *each*. Once understood as a constraint on the mereological structure of distributivity, the extensive-intensive contrast follows. The intensive measurement in (18) measures the size of each water container, which does not change with more girls buying more water. However, the extensive measurement in (18) measures the water volume, which does change with more girls buying more water.

1.3.3 Generalized monotonicity: Mandarin *ge* and other uses of *each*

This chapter takes up Mandarin *ge* and the non-binominal uses of English *each*. *Ge* is a an adverbial distributivity marker in Mandarin. Like binominal *each* and Cantonese *saai*, *ge* imposes peculiar requirements on the distributed share (Lin 1998b, Soh 2005, Lee et al. 2009a, Tsai 2009). On one hand, it patterns like binominal *each* in favoring counting quantifiers and measure phrases with an extensive measure function in its distributed share.

(19) Zhe-xie haizi ge kan-le *(liang-chu) dianyin.
these-CL child GE see-ASP two-CL movie
'The children each saw two movies.' Counting quantifier

(20) Zhe-xie haizi ge zhong 100-bang.
these-CL child GE weigh 100-pounds
'The children are 100 pounds each.' Extensive measurement

(21) ??Zhe-xie yinliao ge re 50-du.
these-CL child GE heat 50-degrees
'The drinks are 50 degrees each.' Intensive measurement

On the other hand, it differs from binominal *each* in that it can be licensed by expressions interpreted as dependent on the distributivity key despite their lack of a measurement component. These expressions include pronouns bound by the distributivity key and adjectives with a sentence-internal reading like *butong* 'different':

(22) Zhe-xie haizi ge kan-le ziji dailai-de shu.
these-CL child GE see-ASP self bring-MOD book

‘The children each read the book they brought.’ Bound pronoun

(23) Zhe-xie haizi ge kan-le butong-de shu.
these-CL child GE see-ASP different-MOD book
‘The children each read different books.’ Sentence-internal *butong* ‘different’

I propose that the monotonic measurement constraint, formulated to account for the licensing requirements of binominal *each*, can be generalized to account for the licensing requirements of Mandarin *ge*. More specifically, I argue that *ge* can make use of both dependent individuals and dependent measurements to form the monotonic mapping that is necessary to satisfy its monotonicity constraint.

A further generalization of the monotonicity constraint to allow events and their thematic dimensions to participate in the relevant mappings allows us to subsume the ‘event differentiation condition’ of non-binominal uses of *each* (Tunstall 1998, Brasoveanu and Dotlačil 2015). The event differentiation condition has been argued to be responsible for the additional inferences (indicated in italics) in sentences with *each*:

(24) a. The girls each walked to the park.
≈ The girls walked *separately* to the park.
b. Each girl walked to the park.
≈ Each girl walked to the park *by herself*.

1.4 Conclusion

In short, I motivate to treat distributivity markers that impose selectional requirements on the distributed share as a natural class. Their requirements can be uniformly understood as constraints on the dependencies arising from distributive quantification. The fact that a glance at three languages reveals a host of distributivity markers bearing such constraints is suggestive of two things. First, distributivity markers often have dual functions, signaling the presence of distributive quantification and imposing constraints on the outcome of distributive quantification (see also Balusu 2005, Henderson 2014, Kuhn 2017). Second, for the constraints to be satiable, distributivity must contribute dependencies that are available to compositional semantics, as argued in Krifka (1996a), van den Berg (1996), Nouwen (2003), Brasoveanu (2008), Henderson (2014), and Kuhn (2017).

2

DYNAMIC PLURAL LOGIC WITH MEASUREMENT

2.1 Overview

In this dissertation, functional dependencies are formalized using sets of assignment functions, as in Dynamic Plural Logic (van den Berg 1996; see also Nouwen 2003 for an assignment-free implementation) and Plural Compositional DRT (Brasoveanu 2007, 2008, 2010, 2013; see also Dotlačil 2010, Henderson 2014 and Kuhn 2017). The core semantic framework is based primarily on van den Berg (1996), with the following enrichments and modifications: (i) subsentential compositionality, as borrowed from PCDRT, (ii) Linkean referential pluralities in the range of assignment functions, also borrowed from PCDRT, and (iii) degrees in the range of assignment functions.

These modifications are empirically motivated. Subsentential compositionality is needed for modeling how chunks of meaning are pieced together. Referential pluralities are added to model expressions whose denotations do not have a well-defined atomic tier, such as *a lot of milk* and *a large amount of sewage* (Chierchia 2010). Since functional dependencies involving these expressions are an important subject of investigation in this dissertation, it is necessary to allow them to be generated in the first place. Degrees are needed to model functional dependencies involving individuals and their measurements, which is also an important subject of investigation in this dissertation. To make it easier to distinguish between the original DPlL and the modified version here, I call the current framework Dynamic Plural Logic with Measurement (DPlLM).

This chapter has two parts. In section 2.2, I lay out the core semantics of DPlLM. Then, in

Section 2.3 I compare the major differences between DPlLM, DPlL, and PCDRT.

2.2 Dynamic Plural Logic with Measurement

2.2.1 Types and models

Like PCDRT, DPlLM is a typed logic. Table 2.1 lists all the primitive types and the objects associated with them:

Name	Type	Variables	Examples
Individuals	e	x, y, z	$\mathsf{a}, \mathsf{b}, \mathsf{a}\oplus\mathsf{b}$
Events	v	ϵ, ϵ'	$\mathsf{eat1}, \mathsf{eat2}$
Degrees	σ	d, d'	$\langle \mathsf{2kg}, \mathsf{weight}, \mathsf{a} \rangle$
Variable assignments	s	$g, h, \ldots$	
Truth value	t	–	1,0

Table 2.1: Basic types in DPlLM

In addition to the individual variables available in most versions of DPlL and PCDRT, DPlLM has also events of type v (see Henderson 2014), and degrees of type σ. Events variables will be used when we talk about the event differentiation requirement of adverbial and determiner *each* in Chapter 3. Degree variables will be used to talk about functional dependencies that hold between individuals and their measurements, as in the investigation of binominal *each*, Cantonese *saai*, and Mandarin *ge*.

Previous studies have argued that degrees have more structure than a simple numerical value (Grosu and Landman 1998, Rett 2008, Scontras 2014). In this dissertation, they are modeled as a *triple*: the first coordinate of this triple stores a degree name (i.e., a point on a scale, such as 6kg), the second coordinate stores a measurement dimension (e.g., length, weight), and the third coordinate stores the individual being measured. The representation of the measurement dimension and the individual being measured is motivated by the need to model the functional dependencies between degree variables and other variables.

Function types are recursively built out of primitive types. Table 2.2 lists some frequently used function types in the dissertation:

Name	Type	Abbreviation	Variables	Examples
Info-state	$s \to t$		$G, H, \ldots$	
Dynamic proposition	$(s \to t) \to (s \to t) \to t$	t	p, q	Ann left.
Dynamic property	$e \to (s \to t) \to (s \to t) \to t$	$e \to \mathrm{t}$	P, P'	pretty, smile
Dynamic relation	$e \to e \to (s \to t) \to (s \to t) \to t$	$e \to e \to \mathrm{t}$	R, R'	kiss, see
Measure function	$e \to d$		m	$\mu_{\mathsf{temp}}, \mu_{\mathsf{vol}}$

Table 2.2: Function types in DPlLM

I work with standard models $\mathcal{M} := \langle D_e, D_v, D_d, D_s, \mathcal{I} \rangle$, where D_e is the domain of individuals, D_v is the domain of events, D_d is the domain of degrees, D_s is the domain of variable assignments and I is the basic interpretation function where $I(R) \subseteq D^n$ for any n-ary relation R. I assume that D_e and D_v are each subject to the axioms of classical extensional mereology; that is, they are equipped with partial orders $\leq_e$ and $\leq_v$ and sum operations $\oplus_e$ and $\oplus_v$ such that each $\oplus_i$ is the least upper bound of its $\leq_i$ (for additional details, see Krifka 1998; Champollion 2017). D_d is assumed to a set of triples storing a degree name, a measurement dimension, and an individual being measured. Degree names are totally ordered points on the relevant scales.

2.2.2 Information states, truth and connectives

Like DPlL and PCDRT, a DPlLM information state is a set of assignments. Interpreting a formula or **dynamic proposition** yields a relation between information states (info-states). This differs minimally from the more well known DPL (Groenendijk and Stokhof 1991), which interprets a formula as a relation between assignments, rather than sets of assignments.

Definition 1 (Information state)

An information state is a set of assignments.

Definition 2 (Assignments)

An assignment g takes a variable and when defined, returns a (possibly plural) individual.[1]

Following a common practice in the literature, an information state is represented as a matrix. The first row has the variables, introduced into the info-state so far. The first column lists the assignments in it . All other cells store values obtained by applying an assignment to a variable.

G	...	x	y	...
g_1	...	a	e	...
g_2	...	b	c⊕d	...
...	...	...	...	...

Figure 2.1: A sample info-state

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). 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 referential plurality, comes from considering a single assignment function that assigns a plurality to a variable. When we talk about cardinality and measurement, we need to be careful and keep these two types of plurality apart. Fortunately, thanks to the collective evaluation of lexical relations, to be defined in Section 2.2.4, we only need to make reference to evaluation-level cardinality in this framework.

In addition, following van den Berg (1996) and Brasoveanu (2010), I incorporate the dummy individual $\star$ in the range of assignment functions. The dummy individual $\star$ is the universal falsifier, i.e., any lexical relation with $\star$ as its argument, like $\mathsf{read}(\mathsf{a}, \star)$, is false (Brasoveanu 2010). The dummy individual is useful in a few ways. First, it can be used to model info-states that don't contain any information, as visualized in Figure 2.2. Such an info-state is sometimes referred to as

[1]Note that d-refs are modeled as variables in this work, as in the traditions of Discourse Representation Theory (DRT, Kamp 1981; Kamp and Reyle 1993), File Change Semantics (FCS, Heim 1982, 1983), Dynamic Predicate Logic (DPL, Groenendijk and Stokhof 1991), and Dynamic Plural Logic (DPlL, van den Berg 1996). In PCDRT, d-refs are modeled as individual concepts that take an assignment and return an individual (i.e., type $s \to e$ for individual d-refs; see Brasoveanu 2008:137–138; see also the 'register'-type in Muskens 1996). The difference in the two types of treatment does not pertain to the main points of this dissertation. I opt for the former for its simplicity.

$\{\emptyset\}$. Relatedly, as pointed out in van den Berg (1996:Ch 2.4), information growth can be modeled by replacing dummy individuals with real individuals.

$G_\star$	x	y	z
$g_\star$	$\star$	$\star$	$\star$

Figure 2.2: The dummy info-state

Second, the dummy individual is useful for defining generalized quantifiers in dynamic plural logic, which we will turn to in Section 2.2.8. Assuming the dummy individual also enables us to keep the logic simple, as we can work with total assignments rather than partial assignments (Brasoveanu 2010).

Given an info-state, we can define ways to project values in this info-state and in its sub-states. When a variable stores a set of nominal values, the corresponding projection function yields a set of values, as shown in (1). We can also project values stored in a variable by taking into consideration values stored in another variable, using the parameterized projection functions in (2) and (3).

Definition 3 (Value projections)

(1) $G(u) := \{g(u) \mid g \in G \ \& \ g(u) \neq \star\}$

(2) $G|_{u=\alpha}(u') := \{g(u') \mid g \in G \ \& \ g(u) = \alpha\}$

(3) $G|_{u \in U}(u') := \{g(u') \mid g \in G \ \& \ g(u) \in U\}$

Parameterized projection functions enable us to compute functional dependencies between two (or more) variables, and hence will be frequently used throughout this dissertation.

When a variable stores a plural degree in an info-state, applying the projection function to it yields the concatenation of the plurality of degrees, i.e., a *single* degree, which is a triple, as shown in (4). This triple is obtained by applying the measurement function stored in the second coordinates of the plural degree $\bigoplus G^{i=2}(d)$ to the plural individual stored in the third coordinates of the plural degree $\bigoplus G^{i=3}(d)$. A projection function not only can be parameterized based on a variable, as in $G|_{x=a}(y)$, but it can also be parameterized based on a particular coordinate of a variable, as in

$G^{i=1}(d)$.[2]

A rich degree ontology is assumed to model the fact that measurement tracks the mereological structure of individuals, which are in turned tracked by sets of assignments in an info-state. Because a degree contains all the necessary ingredients for building a new degree, we can effectively model the dependency among degrees, individuals and assignments.[3] The reader should take the triple structure of a degree to mean that all the information necessary for computing a degree are recorded in discourse, including the individuals, the measurement function and the outcome of the measurement. This assumption is partially shared by studies like Schwarzschild (2006) and Wellwood (2015), who assume that at least the dimension of a measure function is contextually provided.

We can choose to only project the numerical coordinate of a degree variable, using the notation in (5). Like the parameterized projection function for nominal variables, the projection function for degree variables can also be parameterized, as shown in (6).

Definition 4 (Degree projections)

(4) $G(d) := \underbrace{\bigoplus G^{i=2}(d)}_{\text{measure function}} \, (\underbrace{\bigoplus G^{i=3}(d)}_{\text{plural individual}})$, where $G^{i=2}, G^{i=3}$ are the second and third coordinates of d stored in G.

(5) $G^{i=1}(d) :=$ the first coordinate of $\underbrace{\bigoplus G^{i=2}(d)}_{\text{measure function}} \, (\underbrace{\bigoplus G^{i=3}(d)}_{\text{plural individual}})$

(6) $G|^{i=1}_{x=a}(d) :=$ the first coordinate of $\underbrace{\bigoplus G^{i=2}_{x=a}(d)}_{\text{measure function}} \, (\underbrace{\bigoplus G^{i=3}_{x=a}(d)}_{\text{plural individual}})$

For concreteness, let us consider the info-state in Figure 2.3.

$G(d)$ yields a triple of the form in (7), and $G^{i=1}(d)$ yields the degree name in (8). The parameterized version $G|^{i=1}_{x=\mathsf{a}}(d)$ returns (9).

---

[2] I thank Robert Henderson (p.c.) for suggesting that the coordinates in a degree triple can be treated as part of the parameterization of the projection function.

[3] An alternative, but not simpler, approach is to let individuals, degrees (as numerical values), and measure functions live as different variables, perhaps introduced by different parts of a degree expression. Although this approach does not require us to posit a degree as a triple, every time a plural degree name is computed it is still necessary to retrieve the measure function and apply it to the individuals that are being mapped to the plural degree name. In other words, a project function still needs to take two variables to compute a degree (as numerical values). An assignment function taking two coordinates in a variable to construct a degree name is isomorphic to an assignment function taking two variables to construct a degree name.

G	...	x	d	...
g_1	...	a	$\langle 100\text{kg}, \mu_{\text{weight}}, \text{a}\rangle$	...
g_2	...	b	$\langle 80\text{kg}, \mu_{\text{weight}}, \text{b}\rangle$	...
...	...	...	...	...

Figure 2.3: Projecting degrees

(7) $$G(d) := \underbrace{\bigoplus G^{i=2}(d)}_{\text{measure function}} \; (\underbrace{\bigoplus G^{i=3}(d)}_{\text{plural individual}})$$
$$= \langle 180\text{kg}, \mu_{\text{weight}}, \text{a}\oplus\text{b}\rangle$$

(8) $G^{i=1}(d) := 180\text{kg}$

(9) $G|^{i=1}_{x=\text{a}}(d) := 80\text{kg}$

A sentence is modeled as a dynamic proposition, a device for changing context. Feeding an input info-state to a dynamic proposition returns a set of updated info-states. The truth of a dynamic proposition is defined in the expected form, i.e, existential quantification over output info-states, following DPL/DPlL/PCDRT.

Definition 5 (Truth)
A dynamic proposition ϕ is true ($= \mathbb{T}$) with respect to an input info-state G iff there is an output state H when ϕ is fed G.

The truth condition of a dynamic proposition is represented with help of the interpretation function $G[\![\cdot]\!]H$, which is a more iconic notation for the relation of info-states. The definitions of the propositional connectives are provided below. They basically follow the definitions in DPL, with assignments substituted for sets of assignments.

Definition 6 (Connectives)

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

2. $G[\![\phi \vee \psi]\!]H = \mathbb{T}$ iff $G[\![\phi]\!]H = \mathbb{T}$ or $G[\![\psi]\!]H = \mathbb{T}$

3. $G[\![\phi \rightarrow \psi]\!]H = \mathbb{T}$ iff $G = H$ & $\forall K.H[\![\phi]\!]K = \mathbb{T} \Rightarrow \exists J.K[\![\psi]\!]J = \mathbb{T}$

4. $G[\![\neg\phi]\!]H = \mathbb{T}$ iff $G = H$ & $\neg\exists K.G[\![\phi]\!]K = \mathbb{T}$

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. Disjunction is externally dynamic but internally static, which is taken from a later part of Groenendijk and Stokhof (1991) (Definition 53, page 88; see also Stone (1992) and Charlow (2014: Chapter 4.6) for discussions of dynamic disjunction, and Rooth and Partee (1982) for discussions of the wide scope behavior of disjunction). An externally dynamic disjunction is useful for analyzing the interaction between Cantonese *saai* and disjunction (to be discussed in Chapter 4). Implication is internally dynamic but externally static. Likewise, negation is internally dynamic but externally static (see Krahmer and Muskens 1995 and van den Berg 1996 for modified versions of the negation operator that tracks variables introduced in its scope).

2.2.3 Variable introduction and the lack of dependence

In DPlLM, existential quantifiers are responsible for introducing variables. Variable introduction, also known as random assignment, is defined as in 7. It introduces all values in a set D as values of a variable u.

Definition 7 (Variable introduction)

$G[\![\exists u]\!]H = \mathbb{T}$ iff there is a set D such that $H = \{g^{u\rightarrow\alpha} \mid g \in G \;\&\; \alpha \in D\}$

where D is a subset of D_e, D_v or D_d in $\mathcal{M}$

This version of variable introduction differs from both the DPlL version and the PCDRT version. Unlike the PCDRT version but like the DPlL version, it does not introduce dependence between the introduced variable and any existing variable. Unlike the DPlL version but like the PCDRT version, a single assignment is free to assign any value, atomic or not, to a variable. The result is a variable introduction that may introduce plurality but is still dependence-free.[4]

[4]Whether or not assignment functions may range over pluralities and whether or not new variable introduction is dependence-free are two independent design choices of a dynamic plural logic.

An example should help us see how this variable introduction works. Assume a model with a domain D containing the individuals c, d and $\mathsf{c}\oplus\mathsf{d}$. Introducing a variable x with $\exists x$ on an info-state G involves the following steps, which are also illustrated in Figure 2.4.

(10) 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 y in the domain of g and d in the range of g. Output 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 the output 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 to 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.

Since we're bringing referential plurality into the range of assignment functions, it is necessary for us to represent pluralities as well as singletons. I have chosen to represent them as atomic individuals and sum-individuals, following the tradition of Link (1983).[5] [6]

Due to the dependency-free nature of variation introduction in DPlLM, in none of these output info-states does the value of the variable y depend on the value of the variable x or vice versa. To better see this, we need the formal notion of value dependence:

Definition 8 (Value dependence)

y is value-dependent on x in an information state G iff there are $a, b \in G\ x.G|_{x=a}\ y \neq G|_{x=b}\ y$

For value dependence between two variables y and x to hold, a variable (say x) should not always

[5] An alternative is to represent pluralities as sets following the tradition of Scha (1981), Schwarzschild (1996) and van den Berg (1996). My decision is primary based on readability–the plurality resulting from sets of assignments are represented as sets already, and having referential pluralities modeled as sets inside these sets is not very aesthetically appealing.

[6] As suggested in Brasoveanu (2011), a bonus for having referential pluralities in the range of assignment functions is that cover-based distributivity as proposed in Schwarzschild (1996) can be modeled without using covers.

G	y
g_1	a
g_2	b

$\exists x$

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

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

Figure 2.4: Variable introduction in DPlLM

be assigned the same value for different values assigned to y. In other words, the values stored in x is not constant relative to the values in y. Obviously, x is not dependent on y in any of the info-states shown in Figure 2.4. Take H_3 and H_5 as examples:

(11) $H_3|_{x=\mathsf{a}}(y) = \{\mathsf{c}, \mathsf{d}\}$

(12) $H_3|_{x=\mathsf{b}}(y) = \{\mathsf{c}, \mathsf{d}\}$

(13) $H_5|_{x=\mathsf{a}}(y) = \{\mathsf{c}, \mathsf{c}{\oplus}\mathsf{d}\}$

(14) $H_5|_{x=\mathsf{b}}(y) = \{\mathsf{c}, \mathsf{c}{\oplus}\mathsf{d}\}$

In both info-states $H(y)$ yields the same value when x is restricted to a different value. Hence, y is not value-dependent on x in H_3 or H_5.

2.2.4 Lexical relations

All lexical predicates are cumulatively closed by default, following the assumptions in Landman (2000), Kratzer (2007), Brasoveanu (2013) and Champollion (2017). For example, if $x \in$ **boys** and $y \in$ **boy**, then $x \oplus y \in$ **boy**; 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. The lexical relation is defined as follows:

Definition 9 (Lexical relation)

$$G[\![R(t_1,...,t_n)]\!]H = \mathbb{T} \text{ is true iff } G = H \ \& \begin{cases} \left\langle \bigoplus[\![t]\!]^G, ..., \bigoplus[\![t_n]\!]^G \right\rangle \in I(R), \text{ if } t_1...t_n \text{ are variables} \\ \left\langle [\![t]\!]^G, ..., [\![t_n]\!]^G \right\rangle \in I(R), \text{ if } t_1...t_n \text{ are constants} \end{cases}$$

I define $[\![t]\!]^G = G(t)$ if t is a variable, and $[\![t]\!]^G = I(t)$ if t is a constant.

Definition 10 (Discourse-level summation)

$\bigoplus G(u) = \bigoplus\{g(u) : g \in G\}$

Lexical relations are satisfied *collectively*, like DPlL defined in van den Berg (1996). I use a summation operator to bring a set of individuals to an individual when checking for lexical relations. In Brasoveanu's (2008) PCDRT, lexical relations are distributively checked, as a result of his definition of variable introduction. I'll return to this issue in Section 2.3.

2.2.5 Cardinality and measurement

Although there are two types of pluralities in DPlLM, measurement of these pluralities is done at a uniform manner. In particular, it is done by collapsing all evaluation-level pluralities into a domain-level plurality and taking its measurement. This is true of both cardinality measurement (Definition 12) and other types of measurement (Definition 13). Specifically. the cardinality test is on the atomic parts of the individual obtained from discourse-level summation (see also Henderson 2014:53 for a similar way to find the atomic parts of a discourse plurality).

Definition 11 (Cardinality)

$G[\![|u| = d]\!]H = \mathbb{T}$ iff $G = H$ & $|\{u' \mid u' \leq \bigoplus H(u) \ \& \ u' \text{ is an atom}\}| = d$

Measurement other than cardinality is defined in a similar way in Definition 12 (i), only this time it is not necessary to access the atomic parts of a referential plurality (i.e., plurality in the range of a single assignment). We just need to apply a measure function (with a parameterized dimension, such as **height** or **volume**) to the referential plurality and obtain a degree, following Krifka (1998)[7]. Cardinality measurement (Definition 11) can receive a notational variant closer to non-cardinality measurement, as shown in Definition 12 (ii). This notation is used when an emphasis on the representation of a measure function is warranted.

Definition 12 (Measurement)

(i) $G[\![\mu\ u = d]\!]H = \mathbb{T}$ iff $G = H$ & $\mu(\bigoplus H(u)) = d$

(ii) $G[\![\mu_{\mathsf{card}}\ u = d]\!]H = \mathbb{T}$ iff $G = H$ & $|\{u' \mid u' \leq \bigoplus H(u)$ & u' is an atom$\}| = d$

2.2.6 Distributivity and dependence

Like DPlL, functional dependencies are generated via distributivity, which is modeled as a distributivity operator δ, in DPlLM.

Definition 13 (Distributivity)

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

The distributivity 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, DPlLM opens up a door for introducing dependency into info-states.

Let's take a concrete example to see how δ_u works. Suppose we have a formula with an existential expression in the scope of a distributive operator, i.e., $\delta_y(\exists x)$. Regarding the info-state G, $\delta_x(\exists y)$ first splits up the input info-state along the y dimension, resulting in two atomic sub-states, as shown in Figure 3.10. 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, after collecting the intermediate

[7] In fact, 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). Suppose John weights 68 kilograms. This fact is represented as follows. The function weight maps John to a degree, and then the function kilogram maps the degree to the number 68.

sub-states to form the set of output info-state, some of the output info-states actually exhibit value dependence between x and y, for example H_2 and H_3 in this case.

$$
\begin{array}{c|c} G & x \\ \hline\hline g_1 & \mathsf{a} \\ g_2 & \mathsf{b} \end{array}
\overset{\delta_x}{\Rightarrow}
\left(
\begin{array}{cc}
 & \begin{array}{c|c|c} G|_{x=\mathsf{a}} & x & y \\ \hline\hline g_1 & \mathsf{a} & \mathsf{c} \end{array} \\
\begin{array}{c|c} G|_{x=\mathsf{a}} & x \\ \hline\hline g_1 & \mathsf{a} \end{array} \overset{\exists y}{\Rightarrow} & \begin{array}{c|c|c} G|_{x=\mathsf{a}} & x & y \\ \hline\hline g_1 & \mathsf{a} & \mathsf{d} \end{array} \\
 & \begin{array}{c|c|c} G|_{x=\mathsf{a}} & x & y \\ \hline\hline g_1 & \mathsf{a} & \mathsf{c}\oplus\mathsf{d} \end{array} \\
 & \begin{array}{c|c|c} G|_{x=\mathsf{b}} & x & y \\ \hline\hline g_2 & \mathsf{b} & \mathsf{c} \end{array} \\
\begin{array}{c|c} G|_{x=\mathsf{b}} & x \\ \hline\hline g_2 & \mathsf{b} \end{array} \overset{\exists y}{\Rightarrow} & \begin{array}{c|c|c} G|_{x=\mathsf{b}} & x & y \\ \hline\hline g_2 & \mathsf{b} & \mathsf{d} \end{array} \\
 & \begin{array}{c|c|c} G|_{x=\mathsf{b}} & x & y \\ \hline\hline g_2 & \mathsf{b} & \mathsf{c}\oplus\mathsf{d} \end{array}
\end{array}
\right)
\Rightarrow
\begin{array}{c}
\begin{array}{c|c|c} H_1 & x & y \\ \hline\hline h_1 & \mathsf{a} & \mathsf{c} \\ h_2 & \mathsf{b} & \mathsf{c} \end{array} \\[2ex]
\begin{array}{c|c|c} H_2 & x & y \\ \hline\hline h_1 & \mathsf{a} & \mathsf{c} \\ h_2 & \mathsf{b} & \mathsf{d} \end{array} \\[2ex]
\begin{array}{c|c|c} H_3 & x & y \\ \hline\hline h_1 & \mathsf{a} & \mathsf{c}\oplus\mathsf{d} \\ h_2 & \mathsf{b} & \mathsf{c} \end{array} \\[2ex]
\ldots
\end{array}
$$

Figure 2.5: Introducing dependency with help of a distributivity operator

For concreteness, as shown in (15) and (16), 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).

(15) $H_2|_{x=\mathsf{a}}(y) = \{\mathsf{c}\}$

(16) $H_2|_{x=\mathsf{b}}(y) = \{\mathsf{d}\}$

2.2.7 Types of functional dependencies in DPlLM

Now that there is a way to build a nontrivial functional dependency, we can discuss different types of functional dependencies depending on how variables stand in relation to each other. As will be shown in the subsequent chapters, there are distributivity markers that impose constraints on the type of dependencies a distributive quantification brings about.

We have already seen the first type of dependency. Two variables in a dependency are said to exhibit value dependence (Definition 8, repeated below) when one co-varies with the other.

Definition 8 (Value dependence)

y is value-dependent on x in an information state G iff there are $a, b \in G\ x.G|_{x=a}\ y \neq G|_{x=b}\ y$

In addition to value dependence, there are dependencies that exhibit structural dependence. The formal definition of structural dependence for individual variables is given below:

Definition 14 (Structural dependence (for individual variables))

y is structurally dependent on x in an information state G iff (i) and (ii) holds.

i. there are distinct nonempty sets $A, B \subseteq G(x)$: $G|_{x \in A}(y) \neq G|_{x \in B}(y)$

ii. for all distinct nonempty sets $A, B \subseteq G(x)$: if $A \subseteq B$, then $G|_{x \in A}\ y \subseteq G|_{x \in B}\ y$.

(i) is equivalent to value dependence (see Definition 8). (ii) states that when x stores more individuals, y should not store fewer individuals. This essentially requires that the functional dependency between x and y be correlated for their cardinalities. Since variable introduction in DPlLM is dependence-free, a newly introduced variable also does not stand in any structural dependence relation with another variable. Again, take G_3 and G_5 as examples:

(17) $G_3|_{x \in \{\mathsf{a}\}}(y) = \{\mathsf{c}, \mathsf{d}\}$

(18) $G_3|_{x \in \{\mathsf{b}\}}(y) = \{\mathsf{c}, \mathsf{d}\}$

(19) $G_3|_{x \in \{\mathsf{a}, \mathsf{b}\}}(y) = \{\mathsf{c}, \mathsf{d}\}$

(20) $G_5|_{x\in\{\mathsf{a}\}}(y) = \{\mathsf{c}, \mathsf{c}\oplus\mathsf{d}\}$

(21) $G_5|_{x\in\{\mathsf{b}\}}(y) = \{\mathsf{c}, \mathsf{c}\oplus\mathsf{d}\}$

(22) $G_5|_{x\in\{\mathsf{a},\mathsf{b}\}}(y) = \{\mathsf{c}, \mathsf{c}\oplus\mathsf{d}\}$

It is easy to see that clause (i) is violated for the functional dependency between x and y in these info-states. In other words, there is no structural dependence between x and y.

As pointed out earlier, when a dependent variable stores individual values, value dependence entails structural dependence. For this reason, y is also structurally dependent on x in the info-states H_2 and H_3. The following calculation confirms this: at least two different subsets stored in x are linked to different values in y, and more values in x are never linked to fewer values in y.

(23) $H_2|_{x\in\{\mathsf{a}\}}(y) = \{\mathsf{c}\}$

(24) $H_2|_{x\in\{\mathsf{b}\}}(y) = \{\mathsf{d}\}$

(25) $H_2|_{x\in\{\mathsf{a},\mathsf{b}\}}(y) = \{\mathsf{c}, \mathsf{d}\}$

When a degree variable is introduced inside the scope of a distributivity operator, the relevance of structural dependence comes in. To see this, let us consider the following info-state, with d and d' introduced as a result of distributive quantification over values stored in x.

G	x	d	d'
g_1	a	$\langle 100\mathsf{lbs}, \mathsf{weight}, \mathsf{a}\rangle$	$\langle 36.5, \mathsf{temp}, \mathsf{a}\rangle$
g_2	b	$\langle 100\mathsf{lbs}, \mathsf{weight}, \mathsf{b}\rangle$	$\langle 36.5, \mathsf{temp}, \mathsf{b}\rangle$

Figure 2.6: Extensive measurement

Recall that degrees are modeled as triples. The first coordinate of such a triple stores a degree name, modeled as a point on a scale, the second coordinate stores a measure function, and the third coordinate stores the (possibly plural) individual being measured. When the measure function stored in the second coordinate is *extensive*, as in the case of d, it tracks the size of the third coordinate, resulting in a bigger value (i.e., a higher point on a scale) in the first coordinate when there is a bigger plurality in the third coordinate.

(26) $G|^{i=1}_{x\in\{\text{a}\}}(d) = 100\text{lbs}$

(27) $G|^{i=1}_{x\in\{\text{b}\}}(d) = 80\text{lbs}$

(28) $G|^{i=1}_{x\in\{\text{a,b}\}}(d) = 180\text{lbs}$

However, when the measure function stored in the second coordinate is intensive, as in the case of the temperature measure function stored in d', it fails to track the size of the third coordinate. As a result, even when there are more values in the third coordinate, the first coordinate does not necessarily have a bigger value, as demonstrated below:

(29) $G|^{i=1}_{x\in\{\text{a}\}}(d') = 36.5°\text{C}$

(30) $G|^{i=1}_{x\in\{\text{b}\}}(d') = 36.5°\text{C}$

(31) $G|^{i=1}_{x\in\{\text{a,b}\}}(d') = 36.5°\text{C}$

In many cases, measuring the plurality in the third coordinate with an intensive measure function may not be straightforward. For example, the temperature of two non-overlapping drinks cannot be easily determined. Nor can the speed of two separate cars. However, it does not mean that there is no way of measuring a collective temperature (or speed). One way to do so is to calculate an *average* temperature (or speed). Since an average always falls somewhere between two (or more) measurements, it still fails to be additive.

The following definition of structural dependence for degrees captures the above intuition:

Definition 15 (Structural dependence (for degree variables))
d is structurally dependent on x in an information state G iff (i) and (ii) holds.

i. there are distinct nonempty sets $A, B \subseteq G(x)$: $G|^{i=1}_{x\in A}(d) \neq G|^{i=1}_{x\in B}(d)$

ii. for all distinct nonempty sets $A, B \subseteq G(x)$: if $A \subseteq B$, then $G|^{i=1}_{x\in A}(d) \leq G|^{i=1}_{x\in B}(d)$.

2.2.8 Generalized quantifiers

Most of the work in this dissertation does not use a full-blown semantics of dynamic generalized quantifiers but a simplified version commonly used in studies like Brasoveanu (2008), Henderson

(2014), Kuhn (2017). The simplified version introduces one less d-ref (i.e., the one corresponding to the restriction property) and maximization (i.e., the corresponding maximization over the restriction d-ref) than the full-blown one. I include a discussion of dynamic generalized quantifiers because additional adjustments need to be introduced on top of the original formulation of dynamic generalized quantifiers in van den Berg (1996), due to the incorporation of referential pluralities.

Like DPlL and PCDRT, the definition of generalized quantifiers in DPlLM is built on the classical theory of generalized quantifiers (Barwise and Cooper 1981; Keenan and Stavi 1986). In the classical theory, a generalized quantificational determiner (Det) denotes a relation of two sets of individuals (type $e \to t$). In DPlLM, a generalized quantificational determiner ($\mathbf{Det}$) also denotes a relation, of two dynamic propositions (type $(s \to t) \to (s \to t) \to t$). Recall that a dynamic proposition is a relation of two info-states, and a d-ref in an info-state stores a set of values, assigned to it by a set of assignments. Therefore, it is possible to construct a set of individuals from a dynamic proposition with help of d-refs (see also Dekker (1993) on the related operation Existential Disclosure).

A schema for relating static distributive generalized quantifier to their dynamic correlates is given in Definition 16.[8] The notations used come from Brasoveanu (2010) (the translation schema offered in van den Berg 1996 Section 4 of Chapter 4 is less compact but essentially the same, except for the complexity involved in $\mathsf{Det}_{\oplus}$, to be made precise in Definition 19).

Definition 16 (The translation schema of Det)

$$\mathbf{Det}^{u,u'}(\phi,\ \psi) := \mathsf{max}^{u}(\delta_u(\phi)) \wedge \mathsf{max}^{u' \sqsubseteq u}(\delta_{u'}(\psi)) \wedge \mathsf{Det}_{\oplus}(u, u')$$

In simple words, a dynamic determiner introduces two d-refs u and u', and require that they stand in the relationship specified by the static correlate of the dynamic quantificational determiner $\mathsf{Det}_{\oplus}$. The two d-refs are subject to two additional conditions. First, both u and u' have to be maximal, relative to a restriction proposition ϕ and a scope proposition ψ, respectively. Definition 17 spells out how maximization works. In addition, the subset relation between u' and u is structure-preserving, so that the values in u' stand in the same relationship that the corresponding values in u stand with values in other variables. Structure-preserving subset relation is spelled out in detail in Definition

[8]It is possible to define non-distributive quantifiers. I refer readers who are interested in non-distributive quantifiers to van den Berg (1996:Ch. 3, Ch. 4.4)

18.

The maximization operator, defined in Definition 17, introduces a set of entities and the sum of these entities is the maximal one that satisfies ϕ. Restrictor maximization and scope maximization make sure that the values satisfying the restriction formula and the scope formula are maximal values. They are important as we don't want to accidentally make true *every girl left* in a scenario of five girls by only introducing a proper subset of these girls via variable introduction and checking that every single one of these values left.

Definition 17 (Maximization operator)

$G[\![\mathsf{max}^u(\phi)]\!]H = \mathbb{T}$ iff $G[\![\exists u \wedge \phi]\!]H = \mathbb{T}$ and $\neg\exists K. \bigoplus H(u) < \bigoplus K(u)$ & $G[\![\phi]\!]K = \mathbb{T}$

The structure-preserving subset relation is defined in Definition 18, following Brasoveanu (2010) (see also subset assignment in van den Berg 1996, with the symbol $\subseteq$ in place of $\sqsubseteq$). It requires that a d-ref u' inherits all the dependencies established between the corresponding values in u and values in other variables. This can be achieved by forcing all the assignments that assign a value to u' also assign the same value to u. If there is any value in u that is not assigned to u', the relevant assignments assign a dummy individual $\star$ to u'. A sample of values stored in u and u' ($u' \sqsubseteq u$) is given in Figure 2.7. The values stored in u' form a subset of the values stored in u. In addition, except for the assignment that assigns a dummy individual, all the assignments assign the same values to u and u'. The first matrix in Figure 2.7 fulfills the requirements of $u' \sqsubseteq u$, but the other two violate condition (a) and condition (b) in Definition 18.

Definition 18 (Structure-preserving subset relation)

$G[\![u' \sqsubseteq u]\!]H = \mathbb{T}$ iff $G = H$ and

a. $\forall h \in H.h(u') = h(u) \vee h(u') = \star$ and

b. $\forall h \in H.h(u) \in H(u') \rightarrow h(u) = h(u')$

$\mathsf{Det}_{\oplus}$ represents the standard GQ relation of sets of elements. Since these elements may be in the form of pluralities in the range of assignment functions, it is necessary to extract all the atomic parts from these sets of elements.

G	u	$u'(u' \sqsubseteq u)$
g_1	a	a
g_2	a	a
g_3	b	⋆
g_4	c	c

(a) Fulfilling $u' \sqsubseteq u$

G	u	$u'(u' \sqsubseteq u)$
g_1	a	a
g_2	a	b
g_3	b	⋆
g_4	c	c

(b) Ruled out by condition (a)

G	u	$u'(u' \sqsubseteq u)$
g_1	a	a
g_2	a	⋆
g_3	b	⋆
g_4	c	c

(c) Ruled out by condition (b)

Figure 2.7: An illustration of the structure preserving subset relation

Definition 19 (The GQ relation of sum entities)

$G[\![\mathsf{Det}_{\oplus}(u,u')]\!]H = \mathbb{T}$ iff $G = H$ and $\mathsf{Det}(\{a \mid a \leq \bigoplus H(u)\ \&\ \mathsf{atom}(a)\},\ \{a' \mid a' \leq \bigoplus H(u')\ \&\ \mathsf{atom}(a')\})$

Note that every dynamic determiner contains a static $\mathsf{Det}_{\oplus}$. As a result, any static GQ can be translated into a dynamic GQ. For example, the sentence in (32) can be translated as (33), with the first sentence contributing the first three conjuncts and the second sentence the last conjunct.

(32) Every student came in. They sat down.

(33) $\mathsf{max}^{x}\big(\delta_x(\mathsf{stdt}\ x)\big) \wedge \mathsf{max}^{x' \sqsubseteq x}\big(\delta_{x'}(\mathsf{come.in}\ x')\big) \wedge \mathsf{every}_{\oplus}(x,x') \wedge \mathsf{sat.down}(x')$

In the first conjunct, a d-ref x is introduced that stores all the students. In the second conjunct, a d-ref x' is introduced that stores a structured subset of x and is a maximal set of students that came in. In the third conjunct, $\mathsf{every}_{\oplus}$ stands for a static subset relation, i.e., $\subseteq$. It says that the set of the atomic students is a subset of the set of all the atomic students who came in. Since the first sentence in (32) introduces two d-refs x and x' storing all the students and all the students that came in, respectively, it is predicted that both d-refs can be picked up by subsequent pronouns. For example, the plural pronoun *they* in (32) refers to *every student* via one of these d-refs.[9]

[9]In theory, both d-refs are available for anaphora. However, Nouwen (2003) shows that anaphora to the scope d-refs is more readily available than anaphora to the restriction d-ref. In the case of *every*, both d-refs store the same individuals, so it does not matter which d-ref is used for resolving the anaphora with *they*.

2.3 Comparisons with DPlL and PCDRT

As discussed in the last section, DPlLM is a hybrid of DPlL and PCDRT. It borrows an essential assumption from each of DPlL and PCDRT, as shown in Figure 2.8.[10] On one hand, DPlLM follows PCDRT and assumes that there are two types of pluralities, i.e., domain-level pluralities and evaluation-level pluralities. On the other hand, DPlLM patterns like DPlL in that introduction of variables doesn't generate dependency. As a result, DPlLM not only can incorporate measurement into dynamic semantics for plurals, but also maintains that dependence between variables can only be generated via distributivity.

$$\text{DPlL} \begin{cases} \text{only evaluation-level pluralities} \\ \text{no variable dependence resulting from } \exists u \end{cases}$$

$$\text{DPlLM}$$

$$\text{PCDRT} \begin{cases} \text{domain-level and evaluation-level pluralities} \\ \text{variable dependence resulting from } \exists u \end{cases}$$

Figure 2.8: Relation of DPlLM, DPlL and PCDRT

2.3.1 Comparison with DPlL

van den Berg (1996) presents and elaborates the formal properties of DPlL. The basic idea of DPlL is that, instead of enriching our ontology with plural entities, as suggested by Link (1983), we can enrich our logic, by assuming that contexts of evaluation consist of sets of assignments. So, Linkean plural individuals are replaced by evaluation-level plurality. The domain of individuals contain only atomic individuals, no sum-entities. As a consequence, measurement only needs to be defined at the evaluation level. For example, the numeral phrase *three boys*, which is translated as (34), requires the values stored in the variable x to be three atomic boys.

[10] There are various less essential similarities and differences among DPlL, PCDRT, and DPlLM. DPlL does not have sub-sentential compositionality, while both PCDRT and DPlLM do. In DPlL and DPlLM, a d-ref is a variable that serves as an argument to an assignment function, but in PCDRT, a d-ref is a function that takes an assignment function as its argument (see also Muskens 1996).

(34) $\exists x \wedge \textsf{boy}\ x\ \wedge\ x = 3$

In DPlL, variable introduction is defined in the same way as DPlLM (i.e., Definition 7), with the exception that the assignment functions do not range over plurals (as the domain of individuals only contain atoms), as shown in 20.

Definition 20 (Variable introduction (DPlL))

$G[\![\exists x]\!]H = \mathbb{T}$ iff there is a set D s.t. $H = \{g^{\alpha \to x} \mid g \in G \ \&\ \alpha \in D\}$

where D is a subset of D_e, which contains *atomic entities*

Cardinality of a variable is determined by the set of individuals in the range of a plurality of assignment functions, as indicated in Definition 21. Since DPlL does not have referential pluralities, there is no need to break down the individuals in the range of each assignment function to determine the cardinality of a variable. This contrasts with cardinality in DPlLM (Definition 11, repeated below), which does require breaking down the referential pluralities into atoms in the range of each assignment and counting all the derived atoms.

Definition 21 (Cardinality (DPlL))

$G[\![x = n]\!]H = \mathbb{T}$ iff $G = H$ & $|H(x)| = n$

Definition 11 (Cardinality (DPlLM))

$G[\![|u| = d]\!]H = \mathbb{T}$ iff $G = H$ & $|\{u' \mid u' \leq \bigoplus H(u) \ \&\ u'$ is an atom$\}| = d$

Without extra assumptions, DPlL has difficulty modeling mass nouns, like *water* and *coffee*. This is because mass nouns do not have a well-defined atomic tier (e.g., Quine 1960, Link 1983, Chierchia 2010). Hence, $\exists_{\text{DPlL}} x$ cannot assign appropriate values to a variable with the *water* property. This problem could be resolved by assuming that the domain of individuals contain atomic units of a mass noun. However, the defect of this solution is also obvious. It predicts that mass nouns can be counted without requiring the occurrence of any explicit measure units. This prediction does not sit well with facts from natural language about mass nouns and measurement.

If we would like to handle mass nouns in a dynamic semantics for plurals, the easiest way is to assume that the domain of individuals in the model is closed under sum formation. That is, sum entities are available and they can have the property characterized by a mass noun. Adding sum entities to DPlL directly leads to DPlLM, which inherits most definitions from DPlL.

2.3.2 Comparison with PCDRT

DPlLM is similar to PCDRT in the assumption that a single assignment function may range over pluralities. In other words, referential pluralities are allowed in both PCDRT and DPlL. However, DPlLM differs from PCDRT in two important respects: (i) variable introduction and evaluation of lexical relations. Since the design feature of variable introduction determines the design feature of evaluating lexical relations, I first take up variable introduction and then turn to lexical relations.

Variable introduction

An essential difference between DPlLM and PCDRT lies in whether variable introduction is allowed to introduce dependence among variables. In a series of studies (Brasoveanu 2007, 2008, 2010), Brasoveanu defines variable introduction as in Definition 22. A variable introduction ξx is successful as long as each input assignment g has a successor output assignment h that differs from g at most on the value of x and vice-versa, each output assignment h has a predecessor input assignment g that differs at most from h on the value assigned to x. In other words, as long as all the assignments in an input info-state are preserved in the corresponding output info-state, there is no restriction on what value each assignment associates with x. They may associate the same value(s) with it (in which case there is no dependence) or they may associate different values with it (in which case there is dependence). This is crucially different from the DPlLM variable introduction, which forces all assignments to assign the same value(s) to a variable.

Definition 22 (Variable introduction (PCDRT))

$G[\![\xi x]\!]H = \mathbb{T}$ iff for all $g \in G$, there is a $h \in H$ s.t. $h = g^{x\to\alpha}$ and

for all $h \in H$, there is a $g \in G$ s.t. $h = g^{x\to\alpha}$,

where α is an element in D_e

Assume a model with an individual domain including c, d and $\mathsf{c} \oplus \mathsf{d}$. Introducing a variable x with ξx results in many more outputs than with $\exists x$ in DPlLM, as visualized in Figure 2.9.[11] Crucially,

[11] ξx generates all the outputs $\exists x$ generates plus (potentially infinitely) many more.

the output of ξx can give rise to dependencies when α belongs to a non-singleton domain. As demonstrated by the output info-states $G_4 - G_7$ in Figure 2.9, the values stored in y are associated with different values stored in x.

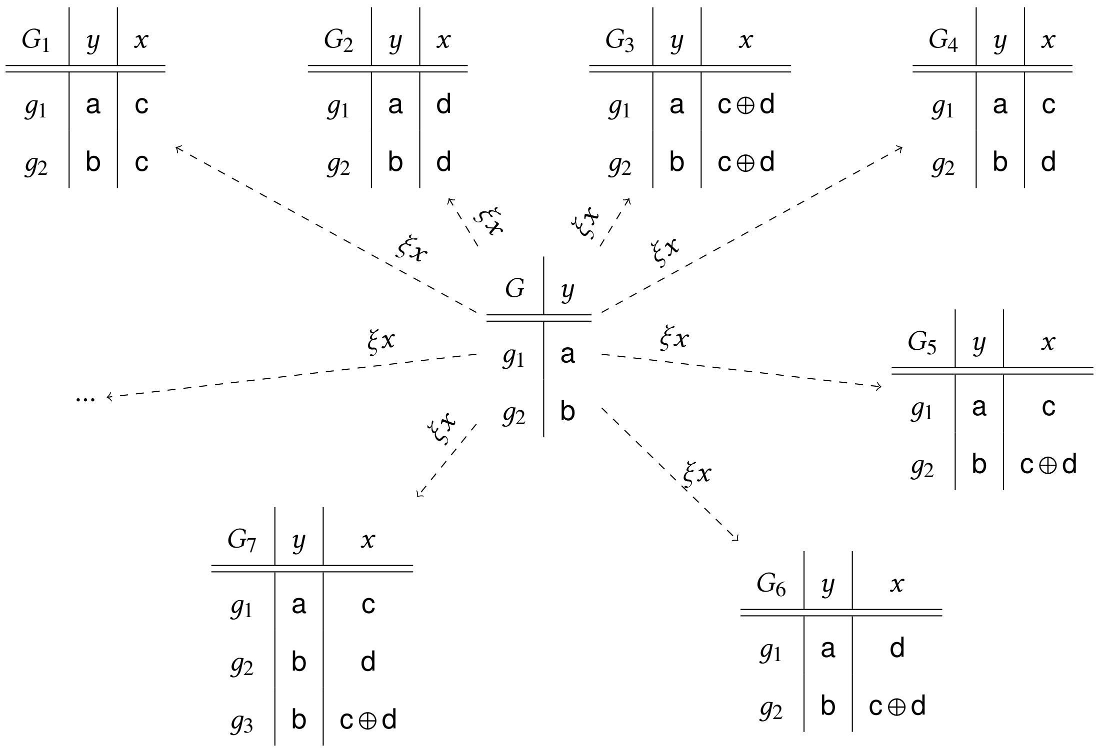


Figure 2.9: Variable introduction in PCDRT

Evaluation of lexical relations

Not only does this version of variable introduction bring in dependencies, it also requires a **distributive evaluation of lexical relations**, as indicated in Definition 23.

Definition 23 (Lexical relations (PCDRT))

$G[\![R(x_1, \ldots x_n)]\!]H = \mathbb{T}$ iff $G = H$ & $\forall h \in H. \langle h(x_1), \ldots, h(x_n)\rangle \in I(R)$

A lexical relation holds between two variables just in case for every assignment, the pair of values associated with the two variables stand in the said relation. This contrasts with how lexical relations are evaluated in DPlLM (as well as DPlL). According to Definition 9 (repeated below), lexical relations are evaluated *collectively* in DPlLM. What this means is that a lexical relation holds between two variables just in case the collective pair of values assigned to the two variables by all the assignments stand in the said relation.

Definition 9 (Lexical relations (DPlLM))

$$G[\![R(t_1,\ldots,t_n)]\!]H = \mathbb{T} \text{ is true iff } G = H \;\&\; \begin{cases} \left\langle \bigoplus[\![t]\!]^G, \ldots, \bigoplus[\![t_n]\!]^G \right\rangle \in I(R), \text{ if } t_1 \ldots t_n \text{ are variables} \\ \left\langle [\![t]\!]^G, \ldots, [\![t_n]\!]^G \right\rangle \in I(R), \text{ if } t_1 \ldots t_n \text{ are constants} \end{cases}$$

The contrast in how lexical relations are evaluated stems from the distinct design features in variable introduction in the two logical systems. In DPlLM, a newly introduced variable does not depend on any other extant variable by default. So, every value in a newly introduced variable is ‘related’ to every value in another variable (for all info-states). To evaluate a lexical relation distributively will result in a very strong claim—every value in a variable stands in the said relation to every value in another variable. To avoid making such a strong claim, it is necessary to check a lexical relation collectively.

The state of affairs is very different in PCDRT, in which a newly introduced variable may depend on an extant one. Since a value in a newly introduced variable does not stand in relation to every value in another variable for all info-states, checking lexical relations distributively has a much milder effect. Some info-states will be ruled out if the two variables do not stand in the right relation, but others will survive. Moreover, since dependencies between variables can be generated, checking lexical relations collectively will ‘waste’ the dependencies. After all, collective evaluation of a lexical relation ignores all the dependencies among the relevant variables.

Which variable introduction to prefer?

The reader may notice that there is a principled way to map a DPlLM-type of logic to a PCDRT-type of logic, by elaborate use of the distributivity operator δ_x. Specifically, by *always* evaluating variable introduction and lexical relations in DPlLM in the scope of a distributivity operator leads us to the same variable introduction and lexical relation evaluation in PCDRT.

The essential difference between PCDRT and DPlL/DPlLM, then, lies in whether distributivity and the dependencies arising from it are taken to come from a default distributive interpretation procedure of the logic or from a lexical source (i.e., an overt or covert distributivity operator that accompanies the distributive interpretation). The former presumably generates dependencies in more contexts than the latter. There is no *a priori* basis to favor one over the other. However, two empirical consideration have been discussed that shed light on the choice, one supporting the PCDRT variable introduction the other supporting the DPlL/DPlLM variable introduction. I discuss them in turn.

The empirical fact that supports PCDRT variable introduction comes from mixed readings of donkey sentences (Brasoveanu 2007, 2008). An example involving a mixed reading is given in (35).

(35) Every person who buys a^{x} book on amazon.com and has a$^{x'}$ credit card uses it$_{x'}$ to pay for it$_{x}$.

Brasoveanu (2008) notes that this sentence is compatible with a situation in which (i) every person buys more than one book from amazon.com, (i) every person has more than one credit card, and (iii) he or she uses different credit cards to pay for the books. His account has two assumptions. First, *a book* introduces a maximal set of books, for each person. This gives rise to the strong reading of *a book*. Second, *a credit card* introduces a non-maximal credit card for *each* book introduced earlier. This models the weak interpretation of *a credit card*. In addition, due to the fact that *a credit card* is introduced in the distributive scope of *a book*, there can be co-variation between the books and the credit cards. Based on this type of dependence between two donkey indefinites, Brasoveanu (2007, 2008) motivates the PCDRT variable introduction, which allows a newly introduced variable to be dependent on an extant one without a distributivity operator.

However, Champollion et al. (to appear) argue that mixed readings of donkey sentences should be analyzed as involving truth value gaps to be adjusted pragmatically. If their proposal is on the right track, then there is no need to use dependency-introducing variable introduction to model mixed readings of donkey sentences.

The empirical consideration that favors DPlLM variable introduction concerns expressions that require dependencies to be licensed. Examples of such expressions are dependent indefinites and the distributivity markers investigated in this dissertation. As many studies have observed, dependent indefinites piggybacks on a distributive interpretation to be licensed (Farkas 1997, 2002b, Yanovich

2005, Brasoveanu and Farkas 2011, Henderson 2014, Kuhn 2017). In this dissertation, I show that the hallmark inferences associated with dependent indefinites can also be borne by distributivity markers. Both classes of phenomena suggest that there is an intimate relationship between distributive interpretation and dependencies. However, this relationship is understood quite differently in DPlLM and PCDRT.

In DPlLM (as well as DPlL), dependencies arise only when variable introduction is interpreted in the scope of a distributivity operator. When a distributivity operator is missing, variable introduction does not introduce dependency. For this reason, expressions that require dependencies naturally require a distributivity operator. Since a distributivity operator is responsible for the distributive interpretation, it is naturally predicted that expressions requiring dependencies are parasitic on the distributive interpretation.

For concreteness, consider sentences such as (36-a) and (37-a) under a non-distributive interpretation, as indicated in (36-b) and (36-b), respectively.[12]

(36) a. Three boys^x made five kites^y.

b. $\exists x \wedge \textsf{boys}\ x \ \wedge\ |x| = 3\ \wedge \exists y\ \wedge \textsf{movies}\ y\ \wedge\ |y| = 5\ \wedge\ \textsf{make}\ y\ x$

(37) a. The girls^x solved the problems^y.

b. $\textsf{max}^x(\textsf{girl} x)\ \wedge\ \textsf{max}^y(\textsf{problem}\ y)\ \wedge\ \textsf{solve}\ y\ x$

Since the two sentences pattern similarly in DPlLM (as well as in PCDRT), I only offer a discussion of (36-a) here. Interpreting *three boys* generates a set of output info-states with a variable x that stores a set of boy pluralities whose collective cardinality is three. Then, interpreting *five kites* against each info-state in the output yields another set of output info-states, each of which has an additional variable y that stores a set of kite pluralities whose collective cardinality is five. Note that because of the lack of a distributivity operator, there is no dependency between x and y. Lastly, the verb contributes a test making sure that the collective values in x and the collective values in y stand in a making relation. An illustration of such an update is given in Figure 2.10.

The non-distributive interpretation is compatible with a collective interpretation as well as a

[12](36-a) does not have maximization over the two variables, so there may be more than three boys and five kites that stand in a making relation in the model. Maximization may be introduced but is not relevant to the discussion here.

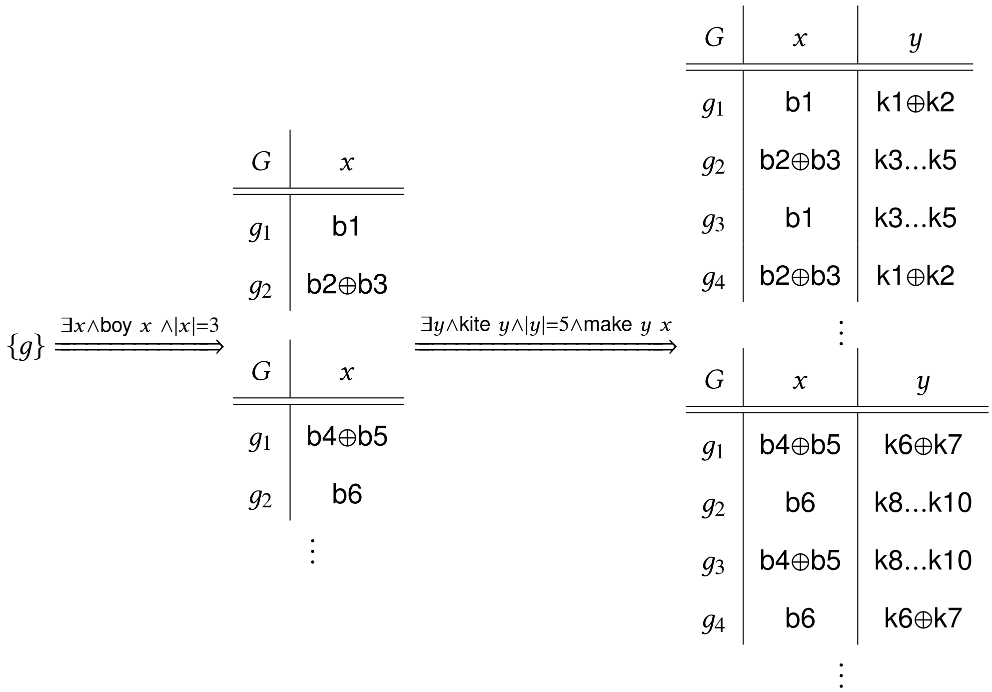


Figure 2.10: No dependency in a collective/cumulative reading (DPlLM)

cumulative interpretation. In fact, without extra machinery, DPlLM does not distinguish between a collective interpretation and a cumulative interpretation (see also Roberts 1987, Link 1998, who also do not distinguish between the two readings; cf. Scha 1981, Landman 2000). Since the non-distributive interpretation lacks dependencies, it is not surprising that it fails to license dependent indefinites. Since only the distributive interpretation may generate dependencies, it is also not surprising why inferences that need access to dependencies show up on distributivity markers.

In PCDRT, by contrast, a non-distributive interpretation may still exhibit dependencies among variables. Consider the PCDRT-theoretic translation of (36-a) in (38). I have used small capitals to translate lexical relations to remind us that lexical relations are evaluated distributively in PCDRT, except for the cardinality tests, which are evaluated *collectively* (i.e., globally) (Brasoveanu 2013).

(38) $\xi x \wedge \text{BOY}\ x \wedge |x| = 3 \wedge \xi\ y \wedge \text{KITE}\ y \wedge |y| = 5 \wedge \text{MAKE}\ y\ x$

In (38), variable introduction makes available two variables, one storing a set of boys and the other storing a set of kites. We can conveniently use the collective cardinality measurement defined for DPlLM in Definition 11 to interpret the cardinality tests. The definition tells us that the cardinality of the two variables are three and five, respectively. Lastly, the two variables are required to stand in a making relation.

Definition 11 (Cardinality (DPlLM))

$G[\![|u| = d]\!]H = \mathbb{T}$ iff $G = H$ & $|\{u' \mid u' \leq \bigoplus H(u)$ & u' is an atom$\}| = d$

A sample flow of update is given in Figure 2.11. What is important is that since variable introduction may encode dependency in PCDRT, many info-states in the output set may encode dependence between the variable storing the boys and the variable storing the kites. The outputs in the sample update are two examples.

$\{g\} \xRightarrow{\xi x \wedge \mathrm{BOY}\ x\ \wedge |x|=3}$

G	x
g_1	b1
g_2	b2⊕b3

G	x
g_1	b4⊕b5
g_2	b6

$\vdots$

$\xRightarrow{\xi y \wedge \mathrm{KITE}\ y \wedge |y|=5 \wedge \mathrm{MAKE}\ y\ x}$

G	x	y
g_1	b1	k1⊕k2
g_2	b2⊕b3	k3...k5

$\vdots$

G	x	y
g_1	b4⊕b5	k6⊕k7
g_2	b6	k8...k10

$\vdots$

Figure 2.11: Dependency in a cumulative reading (PCDRT)

If a dependent indefinite requires variable dependence and a cumulative interpretation provides variable dependence, it is natural to expect that a dependent indefinite should be licensed by a cumulative interpretation.[13] However, as pointed out by Kuhn (2017), dependent indefinites are, in fact, not acceptable without a distributive interpretation. In other words, dependent indefinites and

[13] One may suspect that such a dependent indefinite is independently ruled out for its odd interpretation. I disagree. Consider a hypothetical example below.

(i) Three boys looked for five-five books between them.

If the cumulative reading supported the dependent indefinite, the above sentence would have the following meaning: three boys looked for a total of five books, and not all the boys looked for the same books. This is compatible with the cumulative interpretation.

distributivity seem to go hand in hand.

To model the fact that dependent indefinites are ruled out without a distributive interpretation, Kuhn (2017) and Henderson (2014) introduce different assumptions. Kuhn (2017) suggests building distributivity directly into the semantics of dependent indefinites. In particular, a dependent indefinite is assumed to carry two important components. One component expresses the need for variable dependence (known as evaluation plurality, to be discussed in the next paragraph), another expresses the need for distributively evaluating the numeral contribution.[14] While bundling distributivity and a requirement for variable dependence at the lexical level models the intimate relationship between distributivity and dependency-looking expressions, it is not as explanatory. For example, one may wonder why bundling cumulativity and a requirement for variable dependence does not exist.

As an attempt to limit the distribution of dependent indefinite, Henderson (2014) suggests a way to remove unwanted dependencies generated from variable introduction in PCDRT.[15] In particular, he assumes that expressions that introduce variables into an info-state come in (at least) two types: those that allow variable dependence and those that disallow variable dependence. In other words, Henderson (2014)'s PCDRT has two modes of variable introduction: the dependency-free one proposed in DPlLM (and DPlL, barring referential pluralities), as well as the dependency-introducing one proposed in Brasoveanu (2008).[16] This fact may not be immediately obvious since the locus of the duality does not lie in variable introduction but formulated as a pair of cardinality tests, which are given in Definition 24.[17]

[14] Since the scope of such a distributivity operator is limited to the numeral contribution, Kuhn makes a desirable prediction that a dependent indefinite can be conjoined with a collectively interpreted noun phrase.

(i) A diákok két elöételt és egy-egy föételt rendeltek.
The students two appetizers and one-one main-dish ordered 'The students ordered two appetizers, and N main dishes where N is the # of students.' Hungarian (Kuhn 2017:(21))

[15] Henderson (2014) does not explicitly discuss cumulative interpretations, but his suggestions make clear predictions about cumulative interpretations.

[16] To be more accurate, any variable introduction may encode dependencies. However, while some dependencies are allowed to stay, others are immediately removed.

[17] The reader may notice that evaluation cardinality in Henderson is defined in the same way as the cardinality test in van den Berg (1996). This is not surprising. What evaluation cardinality does is to counting the number of values (possibilities with referential pluralities) in the range of a set of assignments without trying to break down the referential pluralities. In other words, if an assignment ranges over a referential plurality, the plurality still only counts as one, as counting does not go into the subparts of the referential plurality. Since DPlL only has atoms in the range of assignment functions, its cardinality test is devised not to look into the subparts of the individuals in the range of a single assignment function. For this reason, an evaluation cardinality test in Henderson (2014) is the same as the cardinality test in van den Berg (1996).

Definition 24 (Evaluation cardinality)

$G[\![x = 1]\!]H = \mathbb{T}$ iff $G = H$ & $|H(x)| = 1$

$G[\![x > 1]\!]H = \mathbb{T}$ iff $G = H$ & $|H(x)| > 1$

$x = 1$ says that there is only a single, unique value in the range of a collection of assignment functions in an info-state. Recall that $H(x)$ is a set. The set may contain atoms or pluralities, but $x = 1$ requires that there be only one element in the set. So, either a single atom, or a single plurality. By contrast, $x > 1$ says that there must be more than one value in the range of a collection of assignment functions. In other words, there must be at least two assignment functions that assign different values to x. Importantly, Henderson (2014) shows that a noun phrase set to be evaluation singular may still give rise to dependency if there is a distributivity operator scoping over it.

An evaluation cardinality test is to be distinguished from a referential cardinality test (translated as $\text{ONE}\ x$, $\text{TWO}\ x$, etc). The latter is just an ordinary predicate in PCDRT and hence is evaluated distributively. So, $\text{THREE}\ x$ is true when there are three atoms in the range of each assignment function applying to x. The definition of 25 is given below for concreteness.

Definition 25 (Domain cardinality)

$G[\![\text{ONE}(x)]\!]H = \mathbb{T}$ iff $G = H$ & $\forall h \in H.|\{x' \mid x' \leq h(x)$ & $\mathsf{atom}(x)\}| = 1$

According to Henderson (2014), plain indefinites are evaluation singular. So, (36-a) is translated as (39) in Henderson's PCDRT.

(39) $\xi x \wedge \text{BOY}\ x\ \wedge x = 1 \wedge \text{THREE}\ x\ \wedge\ \xi\ y\ \wedge\ y = 1\ \wedge\ \text{FIVE}\ y\ \wedge\ \text{KITE}\ y\ \wedge\ y = 5\ \wedge\ \text{MAKE}\ y\ x$

Since both x and y are required to be evaluation singular, interpreting (39) does not lead to a set of info-states with dependencies between x and y anymore. Rather, the info-states in the output will be akin to the ones in Figure 2.12. Although represented differently, the cumulative reading as represented in Figure 2.12 (Henderson-style PCDRT) and the cumulative reading as represented in Figure 2.10 (DPlLM) encode the same, dependency-free, information.

Although by using evaluation singularity as proposed in Henderson (2014) provides a way to explain why non-distributive readings do not generally license dependent indefinites, the nature of

$$\{g\} \xRightarrow{\xi x \wedge \text{BOY } x \ \wedge x=3} \begin{array}{c} \begin{array}{c|c} G & x \\ \hline\hline g_1 & \text{b1}\oplus\text{b2}\oplus\text{b3} \end{array} \\ \\ \begin{array}{c|c} G & x \\ \hline\hline g_1 & \text{b4}\oplus\text{b5}\oplus\text{b6} \end{array} \\ \vdots \end{array} \xRightarrow{\xi y \wedge \text{KITE } y \wedge y=5 \wedge \text{MAKE } y \ x} \begin{array}{c} \begin{array}{c|c|c} G & x & y \\ \hline\hline g_1 & \text{b1}\oplus\text{b2}\oplus\text{b3} & \text{k1...k5} \end{array} \\ \\ \begin{array}{c|c|c} G & x & y \\ \hline\hline g_1 & \text{b4}\oplus\text{b5}\oplus\text{b6} & \text{k6...k10} \end{array} \\ \vdots \end{array}$$

Figure 2.12: No dependency in a cumulative reading (PCDRT as in Henderson 2014)

evaluation singularity does not sit well with PCDRT's spirit, namely, dependencies among variables are freely available. The use of evaluation singularity is precisely to get rid of these dependencies. If we are to assume that evaluation singularity is associated with the majority of expressions that trigger variable introduction, we lose the only essential difference between PCDRT and DPlLM: freely available vs. restricted dependencies. For this reason, I take Henderson (2014)'s revisions of PCDRT as motivations for a logic for like DPlLM for analyzing expressions that require access to dependencies.

2.4 Summary

We have introduced the framework of DPlLM and discussed its relations to DPlL of van den Berg (1996) and PCDRT of Brasoveanu (2008). In the next three chapters, I use this framework to analyze three distributivity markers: English *each*, Cantonese *saai*, and Mandarin *ge*.

3

Binominal *each*

3.1 Introduction

English distributivity marker *each* has (at least) three uses, as shown in (1) – (3).[1] I refer to these uses as adverbial, determiner, and binominal *each*, respectively. They have taken on various names in the literature, which I have included in (1) – (3) below to facilitate cross reference.

(1) Adverbial *each* (aka 'floated' in Choe (1987a))
The girls *each* saw two movies.

(2) Determiner *each* (aka 'prenominal' in Safir and Stowell (1988))
Each girl saw two movies.

(3) Binominal *each* (aka 'anti-quantifier' in Choe (1987a), 'shifted' in Postal (1974), and 'ad-nominal' in Champollion (2016))
The girls saw two movies *each*.

This chapter is devoted to the third, binominal use of *each*. The other two uses of *each* will be take up in Section 5.4 of Chapter 5. The easiest way to spot binominal *each* is

Binominal *each* has many interesting properties. The primary focus of this chapter is the

[1]There is another instance of *each* that is not taken up in this dissertation, namely, *each* as in *The girls saw each other*. I refer the reader to, among others, Heim et al. (1991), Dalrymple et al. (1998), Beck (2000), Dotlačil (2010), Brasoveanu (2011), LaTerza (2014) for previous studies on this use.

morpho-syntactic and interpretive requirements binominal *each* imposes on its host, i.e., the noun phrase immediately preceding it. They include: *counting quantifier requirement* (Safir and Stowell 1988, Sutton 1993, Szabolcsi 2010), the *variation requirement* (Cable 2014, Champollion 2015, Kuhn 2017), and the *extensive measurement requirement* (Zhang 2013).

I argue that these requirements can be accounted for in a unified manner if binominal *each* is taken to contribute a monotonic measurement constraint on the dependency arising from distributive quantification. Specifically, the constraint requires a monotonic mapping from the size of the plurality contributed by the distributivity key to the measurements contributed by the host of binominal *each*.

Although the semantics proposed for binominal *each* aligns it with distributive numeral markers in broad terms, I would like to highlight an important difference between the analyses developed in this study and the analyses pursued by previous studies on distributive numerals. The difference lies in whether a constraint on a functional dependency makes use of the **mereological structure** of the dependency. Previous accounts of distributive numerals do not access the mereological structure of functional dependencies. However, with help from binominal *each*, I show that it is crucial to treat functional dependencies as having a mereological structure.

To build functional dependencies with mereological structure, I use the version of dynamic plural logic developed in the previous chapter. I show that with the DPlLM architecture, the beyond-distributivity properties of different uses of *each* can be understood as a family of monotonicity constraints, targeting the mereological structure of the functional dependencies arising from distributive quantification.

This chapter proceeds as follows. I start the investigation with binominal *each* (Section 3.2 – Section 3.4), whose beyond-distributivity properties have been widely documented. The lessons learned from the investigation of binominal *each* are then extended to determiner and adverbial *each* in Section **??**. To make a case that binominal *each* indeed impose constraints on the internal mereological structure of distributivity dependencies, I first discuss the beyond-distributivity properties of binominal *each* in Section 3.2. Then, I offer an informal generalization, in Section 3.3, that binominal *each* imposes a measurement-based monotonicity constraint on the internal mereological structure of the functional dependencies arising from distributivity. In Section 3.4, I offer a compositional implementation of the monotonicity constraint in the framework of DPlLM. In Section 3.5, I take up the interaction between binominal *each* and negation. In Section **??**, I propose an extension of the monotonicity constraint to determiner and adverbial *each*. In Section 3.6, I compare my proposal on *each* with previous studies.

3.2 The selectional requirements of binominal *each*

In this section, I discuss three properties of binominal *each* that does not directly follow from it being a distributivity marker. These three properties all have to do with the noun phrase that immediately precedes binominal *each* (underlined in (4)).

(4) The girls saw <u>two movies</u> each.

To facilitate the discussion, let me introduce some terminology for referring to different parts of this sentence. The noun phrase that immediately precedes binominal *each* is called a *host* of binominal *each*. Following the terminology established in Chapter 1, the noun phrase being distributively quantified, typically a plural expression occupying the subject position, is called a *distributivity key*. The whole predicate following the distributivity key is called a *distributed share*.

3.2.1 Variation requirement

Safir and Stowell (1988) is the earliest study, as far as I know, to notice the variation requirement of binominal *each*. They observed that in a sentence like (5), 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 (6), the result is unacceptable.

(5) The girls saw two movies each.

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

Safir and Stowell (1988) treat binominal *each* as a polyadic distributivity operator that quantifies over sets provided by two nominals at the same time (hence the name 'binominal'.) In (5), the quantification results in a one-to-one correspondence between girls and movies, such that each girl saw a different set of two movies.

Moltmann (1991) points out that the one-to-one correspondence condition is too strong. She suggests weakening it to a condition of distinct d-refs, noting that distinct d-refs do not necessarily have distinct values. Recent studies that recognize the variation requirement, such as Cable (2014), Champollion (2015), and Kuhn (2017), borrow insights from distributive numerals and model the variation requirement of binominal *each* along the same lines as the variation requirement of distributive numerals.[2]

Generally speaking, a distributive numeral is a numeral phrase with a morphological marker that induces a distributive interpretation of the sentence. In addition, the morphological marker bears an additional component requiring the numeral phrase to contribute a witness that co-varies with the distributivity key. The following sentence from Kaqchikel (cited from Henderson 2014) illustrates a distributed numeral marked by numeral reduplication:

(7) K-onojel x-ø-ki-kano-j *ju-jun* *wuj*.
E3p-all CP-A3s-E3p-search-SS one-RED book
'They looked for one book each.' Kaqchikel

a. Distributivity inference: Each of them looked for a book.

b. Variation inference: More than one book was looked for.

Couched in various frameworks, Farkas (1997, 2002a,b), Balusu (2005) and Henderson (2014) have proposed a plurality condition for capturing the variation requirement of distributive numerals. The plurality condition requires that a distributive numeral must be associated with at least two distinct values after distributivity is evaluated.[3] I review how the variation requirement of distributive numerals is treated in Henderson (2014) in Section 3.6.1. It suffices at this point to know that attempts to extend his treatment of distributive numerals to binominal *each*, such as Champollion (2015) and Kuhn (2017), have been successful in modeling the variation requirement of binnominal *each*.

A few other strategies have been explored to model the variation requirement. In Choe (1987a),

---

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

[3]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.).

the variation requirement is used to signal the obligatory narrow scope of the host of binominal *each*. A binominal *each* is called a ‘anti-quantifier’ because Choe (1987a) takes the variation requirement to indicate that a host of binominal *each* necessarily takes narrow scope, contrary to the scope flexibility of ordinary quantifiers. To see the contrast between a ‘quantifier’ and an ‘anti-quantifier’, consider (8) and (9). (8) is ambiguous between a wide-scope interpretation of *every girl* (8-a) and a narrow-scope interpretation of the quantifier (8-b), relative to the indefinite. Conversely, we can say that the quantifier contributed by the indefinite is ambiguous between a narrow-scope interpretation and a wide-scope interpretation.

(8) Every girl saw a monkey.

a. For every girl, there is a monkey that she saw.

b. There is a monkey such that every girl saw it.

However, the ‘quantifier’ *one monkey each* in (9) lacks the wide scope interpretation. It is only compatible with a narrow-scope interpretation, due to its variation requirement. For this reason, it is called an ‘anti-quantifier’.

(9) The girls saw one monkey each.

a. For every girl, there is a monkey that she saw.

b. #There is a monkey such that every girl saw it.

However, it must be made clear that while narrow scope may give rise to co-variation, it does not guarantee it. To this see, note that (9-a) is compatible with a scenario is which all the girls saw the same monkey. For this reason, the variation requirement cannot be simply restated as a narrow-scope requirement.

Another possibility that has been considered, in Kuhn (2017), is to generate the variation requirement as an implicature. For example, the wide scope indefinite interpretation (9-b) entails the narrow-scope indefinite interpretation (9-a). 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 covariation implicature. This possibility is briefly considered in Henderson (2014) and discussed in more detail in Kuhn (2015: Ch.3.6).

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 license it. However, distributive numerals, when failed to be licensed, are ungrammatical.[4] Using data from American Sign Language, Kuhn argues that it is desirable to analyze distributive numerals and quantifier-internal adjectives 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 distributive numerals and binominal *each*, it cannot be extended to quantifier-internal adjectives.

In short, the variation requirement of binominal *each* has received a few theoretical treatments. While the narrow scope approach (Choe 1987a) and the implicature approach face some difficulties, the plurality approach defended in many extant studies are empirically adequate for treating the variation requirement.

The beyond-distributivity properties of binominal *each*, however, are not limited to the variation requirement. In fact, there are two other requirements of binominal *each* that cannot be accounted for by studies that only target the variation requirement. I discuss these requirements in the next two subsections.

3.2.2 Counting Quantifier Requirement

It is generally agreed that binominal *each* forms a constituent with its host (Burzio 1986, Safir and Stowell 1988). In addition, studies have documented 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, I believe, comes from Sutton (1993). In particular, Sutton (1993) 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).[5] All other noun phrases are rejected. The contrast is illustrated in (10) and (11).[6]

[4] 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.).

[5] 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).

[6] To some speakers, *some* and *few* are better than the rest in (11) (Simon Charlow, p.c.).

(10) The boys 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.

(11) *The boys 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 handle the counting quantifier requirement take it to be *syntactic* in nature. For example, Zimmermann (2002) takes binominal *each* to only compose with an indefinite. However, this analysis over-generates, as many indefinites in (11) cannot host binominal *each*, such as *some movies*, *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)):

(12) %The men saw a jewel each.

Alternatively, Cable (2014) proposes that binominal *each* takes a number term as one of its arguments. This has the effect of ruling out the hostile hosts in (11). The study may even correctly rule in the friendly hosts in (10) should it treat non-numerical counting quantifier determiners (such as *a few* and *many*) as generalized quantifiers over degrees that must undergo quantifier raising: the movement makes available a degree variable, which has the same type as number terms (Kennedy

2015). Cable's account represents a step forward in understanding the counting quantifier constraint: counting quantifiers are special because they have number terms as part of their semantics. However, in the next subsection, I show that the number component in a counting quantifier does not reliably distinguish hostile hosts from friendly hosts. What matters is the measurement component embedded in a counting quantifier.

3.2.3 Extensive Measurement Requirement and Monotonicity

Zhang (2013) notices that it is insufficient even as a description. Concretely, Zhang (2013) observes that the type of measurement also plays a crucial role in constraining what counting quantifiers may host binominal *each*: extensive measurements give rise to friendly hosts but non-extensive measurements give rise to hostile hosts.

It is widely assumed that numeral expressions such as *two students* and *seven feet* have more structure than meets the eye. In addition to the number word and the common noun, they also contain measure functions like cardinality, height, weight, speed, and temperature. According to Lønning (1987), a measure function denotes a mapping between a class of physical objects and a degree scale that preserves a certain empirically given ordering relation, such as "be lighter than" or "be cooler than." Degrees are further mapped to numbers by unit functions like pound or kilogram. Krifka (1989, 1998) classifies measure functions into two types—extensive and non-extensive measure functions. Crucially, these two types of measure functions differ respect to the property *additivity*. More concretely, weight is extensive since for any object, its weight is equal to the weight of all its parts added together; whereas temperature is non-extensive since the temperature of an object is not always equal to adding up the temperature of its parts.

The examples in (13) and (14) demonstrate Zhang's observation that binominal *each* can only be hosted by a noun phrase with an extensive measure function. To rule out the concern that some of the non-intensive measure functions give rise to a more complex structure, as in the case of speed, or a less natural noun phrase, as in the case of purity, a minimal pair using the measure phrase *60 degrees* is offered. In (13-d), *60 degrees* is a measurement of the angles, and in (14-a), the same form is a measurement of the temperature of drinks.

(13) a. The boys read two books each. cardinality

b. The girls walked three miles each. distance

c. The windows are four feet (tall) each. height

d. The angles are 60 degrees each. angle

(14) 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

Zhang (2013) proposes to understand the extensive measurement requirement as follows: the measurement of the distributivity key should be positively correlated with the measurement of the host. I think Zhang's generalization is essentially correct. Other than Zhang (2013) and my attempt in this chapter, I am not aware of any previous study on binominal *each* that has an account for the extensive measurement requirement.

In fact, binominal *each* is not the only natural language item that cares about the distinction between extensive and non-extensive measurement. Schwarzschild (2002, 2006) points out a similar contrast in pseudo-partitives: pseudo-partitives admit extensive measurement, as in (15), but reject non-extensive measurement, as in (16).

(15) a. two pounds of cherries weight

b. thirty liters of water volume

(16) a. *five degrees Celsius of the water in this bottle temperature

b. *five miles an hour of running speed

In addition, Wellwood (2015) observes similar contrasts in comparatives. Both sentences in (17) can express comparisons involving extensive measurement, but neither can express a comparison involving non-extensive measurement. For example, in (17-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.

(17) 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 (18). This condition requires the part-whole structure of the domain of a measure function be preserved in the domain of degrees.

(18) **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 says the following: only monotonic measure functions can be used in measurement constructions like pseudo-partitives or comparatives. Consequently, extensive measure functions, but not non-extensive ones, pass the condition. Consider a portion of coffee, c, and two of its proper parts, c1 and c2. c necessarily measures a greater degree by volume or weight than that of the parts c1 and c2, but c, c1 and c2 typically have the same temperature. If they don't, the temperature of the c is falls somewhere between the temperature of c1 and the temperatures of c2, making temperature non-monotonic.

It is reasonable to assume that constructions with binominal *each* also obey some version of the monotonicity condition. This will have the following effect: to qualify as a host for binominal *each*, a noun phrase must have a measure function, and the measure function must be an extensive measure function to satisfy the monotonicity condition. It is clear that the monotonicity condition straightforwardly accounts for the sensitivity of binominal *each* towards extensive and non-extensive measure functions. In addition, it illuminates the counting quantifier requirement. Counting quantifiers are essetially measure phrases, typically involving the extensive measure function cardinality. By contrast, bare nouns and indefinites do not contribute any measure function, making them unsuitable hosts. Additionally, Schwarzschild (2006) shows that NPs with Q-adjectives like *many*, *a few*, *a little* and *a lot of* have a syntax similar to measurement phrases and must be associated with a monotonic measure function. In this respect, NPs formed out of them are no different from counting quantifiers.

3.3 A monotonicity constraint for binominal *each*, informally

In this work, I propose that a sentence with binominal *each* has a two-part contribution: distributivity and monotonicity. The former may be contributed by binominal *each* itself, as argued in Kuhn (2017) and many other studies that simply treat *each* as a distributivity operator (e.g., Zimmermann 2002, Dotlačil 2012, Champollion 2017), or by a separate distributivity operator, as suggested in Champollion (2015) following Henderson's (2014) semantics for distributive numerals. I left both options as open possibilities since there is considerable inter-speaker variation regarding how acceptable binominal *each* is when a distributive quantifier is present:

(19) %Every boy saw two movies each.

The monotonicity inference is assembled with help from the distributivity inference as well as the ingredients provided by the host. Leaving a fully compositional implementation until Section 3.4, let me spell out the formation of the monotonicity inference in plain English below.

The distributivity inference provides a set of functional dependencies indicating the relationship between the individual parts of the distributivity key and information provided by various expressions in the share. For example, (20) provides us with, at the very least, a set of functional dependencies encoding the relationships between the boys, the movie-watching events, and the movies being watched. I have singled out the dependency between the boys the movies being watched in Figure 3.1.

(20) The boys saw two movies each.

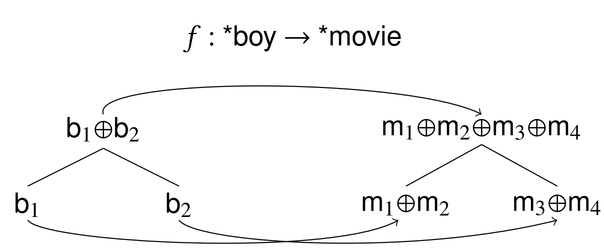


Figure 3.1: Dependency established via distributivity

Here, boy1 saw movie1 and movie2, while boy2 saw movie3 and movie4. The movies seen between the two boys are movie1, movie2, movie3 and movie4. Let's assume a function f that maps each boy to the movies he saw and also sums of boys 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).

The host of binominal *each* provides two important ingredients: (i) a measurement function μ, and (ii) a label for the range of f. The second is important because more than one functional dependency can be formed out of any distributive quantification. The host indicates which functional dependency is being considered.

With f and μ in hand, we can define a monotonic measurement condition checked in association with the functional dependency of distributivity, as in (21) (I abbreviate this condition **dm**, with 'd' a nemonic for distributivity and 'm' a nemonic for monotonicity and measurement).

(21) **Monotonic measurement in association with distributivity (dm, with f)**

A measure function μ satisfies **dm** iff there is a function f such that[7]

a. NON-DECREASING MAPPING

For all $a, a' \in \mathsf{Dom}(f)$. $a \leq a' \rightarrow \mu(f a) \leqslant \mu(f\ a')$, and

b. NON-CONSTANT MAPPING

There distinct $b, b' \in \mathsf{Dom}(f)$. $\mu(f\ b) \neq \mu(f\ b')$

(21-a) requires that the part-whole relation found in the domain of f (i.e., the boys) be mapped **non-decreasingly** to the measurement of the range of f (i.e., the movies). Since f encodes the functional dependency induced by distributivity, this amounts to making reference to the structure of distributivity. In other words, we are referring to parts of a distributivity dependency that stand in a part-whole relation. Modulo the association with f, (21-a) is a standard definition of non-decreasing monotone functions. It is weaker than the definition of monotonic measure functions found in Wellwood (2015), which picks out strictly increasing functions among the non-decreasing ones. I will return to this difference after demonstrating how the definition in (21) works as a whole.

[7] Recall that degrees are modeled as triples in this study and their first coordinates (i.e., degree names) are fully ordered relative to a scale and a relation. So, $\mu(f a) \leq \mu(f a')$ iff the first coordinate of the triple arising from $\mu(f a)$ is at least as great as the first coordinate of the triple arising from $\mu(f a')$, likewise for $\mu(f a) = \mu(f a')$.

(21-b) requires measurement variability in the range of f.

3.3.1 Capturing the extensive measuring requirement

Let me start by demonstrating how the monotonic measurement condition captures the extensive measurement requirement. Consider (20) with a function f as illustrated in Figure 3.1. The measure function in this case is cardinality (or $\mu_{\textsf{card}}$). It is clear that (20) in this setup satisfies (21). First, suppose we take elements b1, b2 and b1⊕b2, the former two are proper subparts of the last one. f maps b1 to m1⊕m2, b2 to m3⊕m4 and b1⊕b2 to m1⊕m2⊕m3⊕m4. The cardinality function μ maps f(b1) to 2, f(b2) also to 2, and f(b1⊕b2) to 4, as shown in Figure 3.2. Since the measurement of the range of f does not decrease (in fact, it increases) as we consider increasingly bigger elements in the domain of f, we can conclude that (21-a) is satisfied. In addition, there are at least two elements in the domain of f that get mapped to elements in the range of f that also yield different measurements. For example, b1 and b1⊕b2 are such a pair, so are b2 and b1⊕b2. We can conclude that (21-b) is also satisfied.

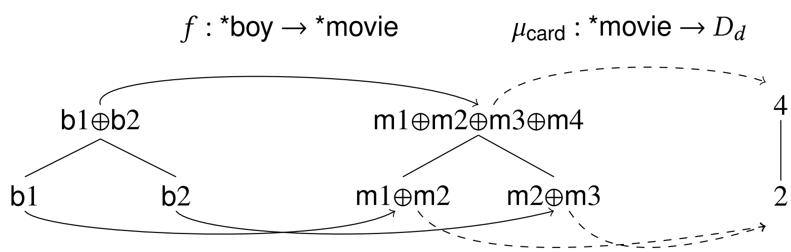


Figure 3.2: Extensive measurement tracks the internal structure of distributivity

(21-a) alone is a rather weak condition. In fact, as long as the measure function involved in the host of binominal *each* is extensive, it is always satisfied, regardless of how many elements there are in the domain and the range of f. One can verify this by constructing scenarios with only one element in the domain of f and/or only one element in the range. In addition, if $\mu_{\textsf{dim}}$ is non-extensive, as long as the range of f is a singleton, or the range of $\mu_{\textsf{dim}}$ is a singleton, (21-a) is satisfied. Therefore, to have the right strength, (21-a) has to be complemented by (21-b).

What (21-b) requires is that the values stored in the range of f must yield different degrees after being measured by a measure function. This rules out the possibility of all values in the range of f having the same measured degree. For example, (22) has a non-extensive measure function temperature (or μ_{temp}) that typically yields a uniform degree for all the values in the range of f, as illustrated in Figure 3.3. It is predicted to fail non-constant mapping, i.e., (21-b), and hence violate the monotonic measurement condition. Note that it does not violate NON-DECREASING MAPPING in (21-a), as the measurement is indeed a non-decreasing mapping of the domain of f, albeit in a trivial way as there is only one degree in the range of μ_{temp}.

(22) *The boys bought 60-degree coffee each.

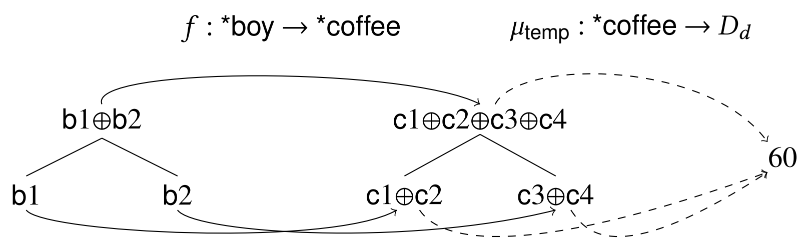


Figure 3.3: Non-extensive measurement dose not track the internal structure of distributivity

If measuring the range of f indeed yields different degrees, as in the case of cardinality measurement as illustrated in Figure 3.2, non-constant mapping is satisfied.

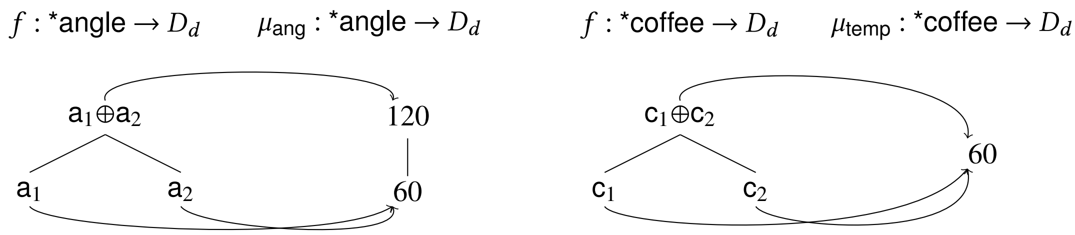
Interestingly, f and μ_{dim} may happen to be the same function, and the contrast between extensive and intensive measurement still holds, as shown in (23-a) and (23-b).

(23) a. *The coffees are 60 degrees each.

b. The angles are 60 degrees each.

In these two examples, a predicative measure phrase helps map individuals in the distributivity key to the corresponding degrees of measurement as indicated by the measure phrase. For concreteness, the coffees (or angles) are distributively checked for their temperature (or degree). So, f encodes a functional dependency between coffees (or angles) and their temperatures (or degrees). μ_{temp} (or μ_{ang}) is identical to f in being a temperature (or angle) measure function. Both sentences satisfy NON-DECREASING MAPPING as stated in (21-a). However, (23-a) fails NON-CONSTANT MAPPING while (23-b) satisfies it. This is because there is only one temperature, i.e., 60 degrees, in association with f in (23-a), but two degrees, i.e., 60 degrees and 120 degrees, in association with f in (23-b).

The contrast is illustrated in Figure 3.4.



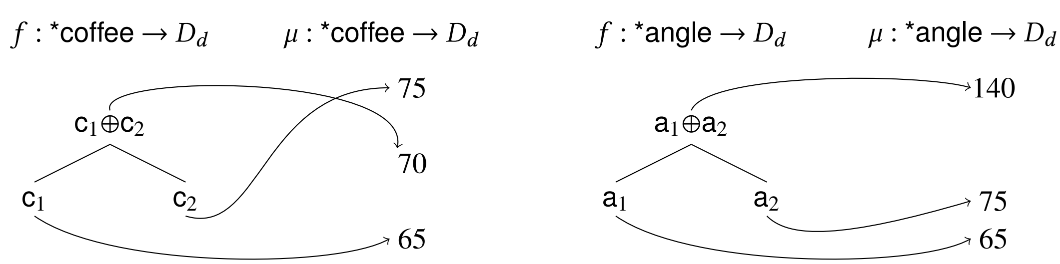
(a) Angle **(b)** Temperature

Figure 3.4: Measure phrase host

One may suspect that (21-b) alone is sufficient to guarantee the variation inference and the privilege of extensive measure functions. It is not. It can be satisfied with a non-extensive 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 (24-a) and (24-b). Figure 3.5 illustrates how f and μ_{dim} works in these two sentences.

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

b. The angles are more than 60 degrees each.



(a) (24-a) **(b)** (24-b)

Figure 3.5: Violation (a) and observation (b) of non-decreasing mapping

(24-a) satisfies NON-CONSTANT MAPPING (as well as evaluation-level plurality), 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 3.5a. By contrast, (24-b) satisfies both NON-CONSTANT MAPPING (as well as evaluation-level plurality) and NON-DECREASING MAPPING, as indicated in Figure 3.5b.

3.3.2 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 at this point. However, compatibility with unit functions like *pound(s)* and *mile(s)* seems to be a rather reliable test: if a determiner-like expression is compatible with measure units like *pounds* and *miles*, then it can form a noun phrase that can host binominal *each*. Some examples are given in Table 3.1.[8]

It has been pointed out 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 them equally well (Safir and Stowell 1988, Szabolcsi 2010, Milačić et al. 2015), as illustrated in (25).

(25) a. %The boys read a book each.

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

Interestingly, *a* is also compatible with unit functions in ways that *some* is not. Of course, 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 (weak) numeral *one* and an existential determiner, while *some* is only an existential quantificational determiner without a measure function component.[9]

(26) a. a mile, a pound, an inch

[8] *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.

[9] It is perhaps too simple to think that *some NPs* have a simple existential quantifier status. For example, it has been observed that when the common noun restriction is a singular count noun, as in *some girl*, the quantifier carries an epistemic effect (Farkas 2002b, Alonso Ovalle and Menendez Benito 2003.

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

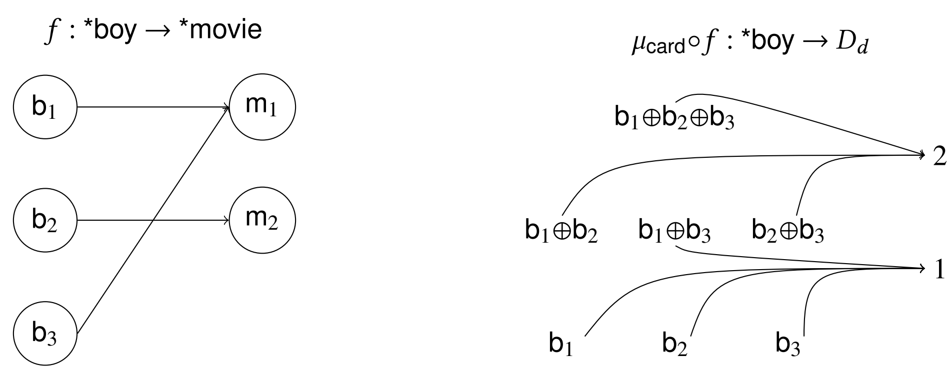
Table 3.1: Expressions that can (and cannot) form a host for binominal *each*

b. *some mile(s), *some pound(s), *some inch(es)

3.3.3 NON-DECREASING + NON-CONSTANT VS. STRICTLY INCREASING

The decision on a weaker form of monotonicity, one in terms of a non-decreasing mapping, instead of a strong form requiring a strictly increasing mapping, as suggested in Wellwood (2015), is empirically motivated. Consider (27). It is judged true in a scenario like Figure 3.6a, in which both boy1 and boy3 saw movie1, while boy2 saw movie2. Since the range of the mapping function serves as the domain of the measure function, we can compose the two functions to form a composite function, $\mu_{\textsf{dim}} \circ f$ as illustrated in Figure 3.6b: the domain of the function is the values associated with the distributivity key, i.e., the boys in this case, and its range is the measured degree of the values introduced by the host, i.e., the cardinality of the movies.

(27) The boys watched one movie each.



(a) Dependency between boys and movies

(b) Dependency established by $\mu_{\textsf{card}} \circ f$

Figure 3.6: Non-decreasing mapping

In this situation, the cardinality of f b_1 is the same as that of f $b_1 \oplus b_3$. Similarly, the respective cardinality of f $b_1 \oplus b_2$ and f $b_2 \oplus b_3$ is the same as that of f $b_1 \oplus b_2 \oplus b_3$. In other words, the composite function is non-injective.

If the monotonicity constraint is formulated to require a strictly increasing mapping, like (28), the situation in Figure 3.6 is predicted to be incompatible with (27), precisely because of the non-injective nature of the composite function illustrated in Figure 3.6b.

(28) **Strictly increasing dm** (rejected)

For all $a, a' \in \mathsf{Dom}(f)$. $a < a' \rightarrow \mu(f\ a) < \mu(f\ a')$

However, formulating the monotonicity constraint as a non-decreasing and non-constant mapping, as in (21), does not run into this problem. The composite function in Figure 3.6b is non-decreasing and non-constant. Therefore, it is predicted that (27) is acceptable in the scenario depicted in Figure 3.6a.

3.4 Formalizing the monotonic measurement condition

Now that the monotonic measurement condition has been established, we are ready to supplement it with a more compositional semantics. In fact, it is not difficult to imagine what kind of framework we need to implement this condition compositionally. The framework should satisfy the following criteria:

- Criterion 1: It should allow us to talk about measure functions of various sorts.
- Criterion 2: It should allow us to represent the functional dependencies arising from distributive quantification and refer back to them. In other words, it should make concrete how f is assembled.
- Criterion 3: Since measurement kicks in after f is established, we need a way to split up the contribution of a host of binominal *each*, evaluating one part (i.e., the basic semantics of the host) 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.

A well-known framework satisfying Criterion 2 is Dynamic Plural Logic of van den Berg (1996) and its close cousin Plural Compositional DRT, devised in Brasoveanu (2007, 2008, 2013).[10] Both approaches have been used to model phenomena that need access to the functional dependencies of distributivity, such as quantificational subordination (van den Berg 1996, Nouwen 2003), quantifier-internal adjectives and reciprocals (Dotlačil 2010), as well as distributive numerals (Henderson 2014, Champollion 2015 and Kuhn 2017).

In Chapter 2, I have developed a hybrid approach, DPlLM, which is intermediate between DPlL and PCDRT, and enriched to include various sorts of measurement. The logic lets assignment functions range over not only atomic individuals, as in van den Berg (1996), but also plural individuals, as suggested in Brasoveanu (2008). However, it sides with van den Berg (1996) in inhibiting dependency introduced by random assignment. To introduce dependency into discourse, a distributivity operator has to be used. This is crucially different from the PCDRT tradition, which allow any random assignment to introduce dependency into discourse. I show, in Section 3.6.1, that this choice explains why the monotonicity condition is only seen with distributive predication.[11]

Criterion 3 essentially asks for a split-scope mechanism. Several alternatives have been explored in the literature. An option is by means of a post-supposition (Henderson 2014, Champollion 2015. Kuhn (2017) points out that post-suppositions, without further assumptions, predict the lack of locality in the licensing of distributive numerals. 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):

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

(30) 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))

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

[10] There are other frameworks that track distributivity dependency. For example, Schein (1993), Lasersohn (1995) and Champollion (2017) develop accounts for distributivity based on event semantics, in which the dependency is retrievable from events. Huang (1996) develops a semantics for distributivity based on skolem functions, in which the distributivity dependency can be retrieved by using skolem functions. The merit of DPlL 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.

[11] Besides van den Berg (1996), many studies have observed that distributive quantification has a much easier time introducing dependency than non-distributive quantification, such as cumulative and collective quantification. Some examples are Nouwen (2003), Solomon (2011), and Bumford (2015).

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

In (29) and (31), the distributive numerals 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 (30) and (32), the distributive numerals are inside ECM clauses, which have been observed not to be a scope island for quantifiers (e.g., May 1985). The fact that distributive numerals 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 distributive numerals.

To model the island sensitivity of distributive numerals, Kuhn (2017) suggests a scope-taking analysis, in which a distributive numeral 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 distributive numerals with a bound pronoun inside them.

(33) 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. While deferring a more detailed discussion until Section 3.4.3, it suffices to note that 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 adopt the higher-order meaning approach for it has better empirical coverage, although the choice is largely immaterial to the main claim that binominal *each* makes reference the mereological structure of a distributivity dependency.

It should be clear by now what kind of framework is needed to account for the novel properties of distributive numerals 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 3.4.1, followed by translating the monotonic measurement condition into this framework, and lastly in Section 3.4.3 the monotonicity condition is implemented in a compositional manner.

3.4.1 Formal background: DPlLM

The background for the account is DPlLM, as outlined in Chapter 2. Recall that in DPlLM, interpreting a formula yields a relation between information states, just as in its cousin logic DPlL and PCDRT. An information state is a set of assignment functions, which is capable of encoding functional dependencies. In addition, by drawing subsets from a set of assignments, we can access the internal mereological structure of the functional dependencies contributed by distributivity.

3.4.2 Monotonic measurement condition in DPlLM

Recall that in Section 3.3, I have sketched the main proposal of this chapter: binominal *each* introduces a constraint known as the monotonic measurement constraint, checking the monotonic property of measure functions relative to the internal mereological structure of the functional dependency established via distributivity.

(34) **Monotonic measurement in association with distributivity (dm, with f)**

A measure function μ is dm iff there is a function f such that

a. NON-DECREASING MAPPING

For all $a, a' \in \mathsf{Dom}(f)$. $a \leq a' \rightarrow \mu(f a) \leqslant \mu(f\ a')$, and

b. NON-CONSTANT MAPPING

There are distinct $b, b' \in \mathsf{Dom}(f)$. $\mu(f\ b) \neq \mu(f\ b')$

The checking of dm is facilitated by a function f that maps values stored in the distributivity key to corresponding values stored in the host. The natural correlate of this f in DPlLM is sets of assignment functions, i.e., info-states. To see this, recall that info-states encode not just values assigned to variables and dependencies among different variables, but also internal structures of these dependencies. In more concrete terms, with help of info-states, not only can we retrieve values associated with the distributivity key and the host of binominal *each*, given that distributivity is externally dynamic in this logic, we can also make precise reference to the corresponding values in the host for all the atomic values and their combinations (i.e., pluralities) in the distributivity key. Having access to this structured dependency allows us to conduct measurement on it to check dm. Translating dm as a dynamic proposition into DPlLM, we obtain (35).[12]

Definition 26 (Monotonic measurement in association with distributivity (dm, in DPlLM))

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

a. $H = G$

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

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

To begin with, dm bears two anaphoric indices. The first one corresponds to the variable introduced by interpreting the distributivity key and the second one corresponds to the variable introduced by interpreting the host. There is a longstanding tradition in granting binominal *each* an anaphoric component, started in the early work of Burzio (1986) and Safir and Stowell (1988) and later adopted in Dotlačil (2012), Cable (2014), and Kuhn (2017). I provide independent justification for using anaphoric indices in Section 3.5, where I discuss how negation interrupts dynamic binding in binominal *each* constructions.

To check for dm of a measure function in DPlLM, we need to access the values stored in the variable the measure function applies to. (35) says that the measure function μ is monotonic on the dependency between x and y iff

- (35-a): Checking dm does not change the info-state in any way (i.e., it's a test).
- (35-b): 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.
- (35-c): In the input info-state, there are at least two sub-parts storing different x''s values that also yield different measurement of y's values.

[12] The domain and the range of the function f in (34) are closed under sum formation. To model this, we consider in an info-state subsets of values assigned to the variable x corresponding to the distributivity key (i.e., the domain of f) and to the variable y corresponding to the host of binominal *each* (i.e., the range of f). Since the host is subject to a measurement transformation, sets of values in y are mapped to mereological sums using $\bigoplus$. See Definition 10 in Chapter 2.

In addition, I propose that the monotonicity condition in (35) is introduced as an 'output context constraint' in the sense of Farkas (2002b) and Lauer (2009, 2012). In particular, (35) 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. As a result, the sentence is *undefined*, rather than *false*. This is to model the fact that sentences with binominal *each* that fail dm (for various reasons) are judged unacceptable and not false, as illustrated below:

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

The constraint is formulated in (37). The connective $\overline{\wedge}$ indicates that the constraint ψ applies after evaluating the at-issue content ϕ.

Definition Output context constraint

(37) $G[\![\phi \mathbin{\overline{\wedge}} \psi]\!]H = G[\![\phi]\!]H$ if $H[\![\psi]\!]H = \mathbb{T}$; otherwise, undefined.

This definition says: the at-issue content given by ϕ has a truth value only if the output context of ϕ 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. This way of understanding the monotonic measurement constraint makes novel and supported predictions about how it interacts with negation and downward monotone quantifiers. These predictions are discussed in Section 3.5.

3.4.3 Composition

Like Nouwen (2003) and Brasoveanu (2008), I assume that DPlLM is a typed logic. It includes basic types and derived types as in (38): e for entities, t for truth values, s for assignments, d for degrees, and a derived type $\tau \to \tau$ for functions.

(38) $\tau ::= e \mid t \mid s \mid d \mid n \mid \tau \to \tau$

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

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

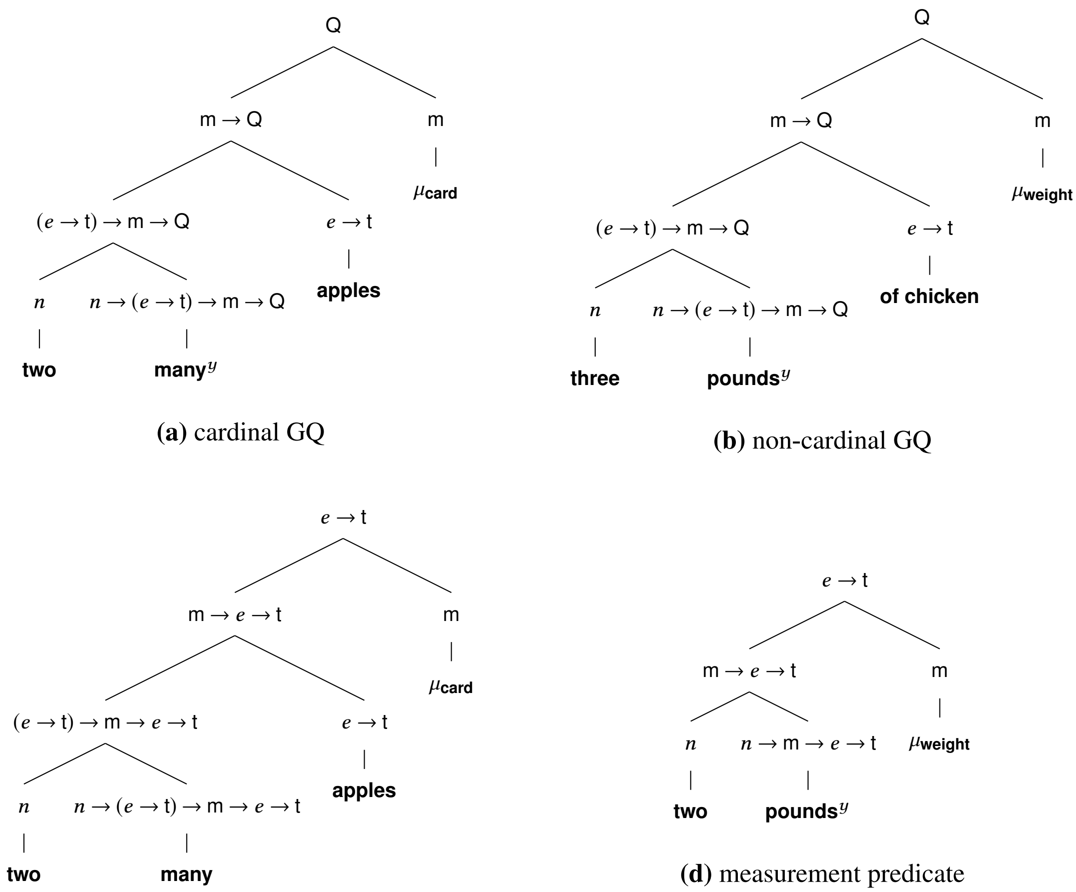
Table 3.2: Type abbreviations

I propose that a noun phrase hosting a binominal *each* is a measure phrase. Depending on whether the measure phrase occurs in an argument position, as in (39-a) and (39-b), or a predicate position, as in (39-c), it has slightly different types.

(39) 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 $(e \to \mathsf{t}) \to \mathsf{t}$. In a predicate position, a measure phrase is simply a predicate, of type $e \to \mathsf{t}$. 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 3.7.

Argumental measure phrases, analyzed as GQs, are shown in Figure 3.7a and Figure 3.7b. 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 number, a property and a measure function and returns a GQ. This measure head is defined in (40-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 number, a property, and a measure function and returns a GQ, as defined in (40-b).



(a) cardinal GQ

(b) non-cardinal GQ

(c) cardinal predicate

(d) measurement predicate

Figure 3.7: Argumental and predicative measure phrases

(40) a. $\textsf{\textbf{many}}^y := \lambda n \lambda P \lambda m \lambda P'.\exists y \wedge P\ y \wedge P'\ y \wedge m\ y = \langle n, m, y\rangle$

b. $\textsf{\textbf{pound}}^y := \lambda n \lambda P \lambda m \lambda P'.\exists y \wedge P\ y \wedge P'\ y \wedge m\ y = \langle n\ \mathsf{lbs}, m, y\rangle$

The cardinality measure head **many** selects (with help of agreement or some other means) a cardinality measure function $\mu_{\textsf{card}}$ (type $e \to n$), while a non-cardinality measure head like **pound** selects an non-cardinality measure function, like $\mu_{\textsf{weight}}$ (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 a predicative and has a nominal predicate (as in *this is two pounds of chicken*), then the measure head takes the same ingredients, but return a predicate rather than a GQ. The corresponding definitions of the predicative measure heads are given in (41-a) and (42-a). Lastly, sometimes a measure phrase may not contain a common head 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 measure phrase. Sample definitions of the measure heads are given in (41-b) and (42-b).

(41) a. $\textbf{many}^{NP} := \lambda n \lambda P \lambda m \lambda u.P\ u\ \wedge\ m\ u = \langle n, m, u\rangle$

b. $\textbf{many}^{MP} := \lambda n \lambda m \lambda u.m\ u = \langle n, m, u\rangle$

(42) a. $\textbf{pound}^{NP} := \lambda n \lambda P \lambda m \lambda u.P\ u\ \wedge\ m\ y = \langle n\ \mathsf{lbs}, m, u\rangle$

b. $\textbf{pound}^{MP} := \lambda n \lambda m \lambda u.m\ y = \langle n\ \mathsf{lbs}, m, u\rangle$

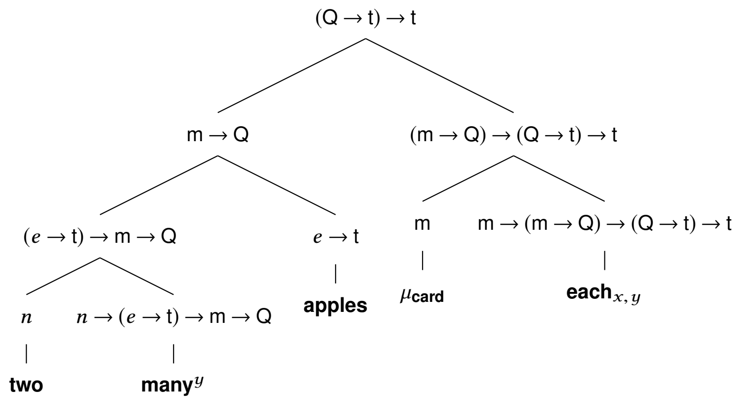
With the assumptions about the internal structure of a measure phrase fleshed out, we are now ready to add binominal *each* to the structure. I assume that binominal *each* attaches to a measure function and turns the whole measure phrase into a higher-order meaning. Concretely, in a cardinal GQ, binominal *each* maps the GQ into a higher-order GQ by turning the measure function from an argument status (it is sought by a $\mathsf{m} \to \mathsf{Q}$ function) to a function status (it now seeks a $\mathsf{m} \to \mathsf{Q}$ function), as shown in Figure 3.8a. 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 3.8b.

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 (43-a), where f may range over any type α. In addition, when a measure phrase does not introduce any discourse variables, as in the case of a predicative measure phrase, *each* only needs to bear one anaphoric index, i.e., the anaphoric index for the variable storing the individuals measured by the measure function $\mu_{\textbf{dim}}$. This is shown in (43-b).

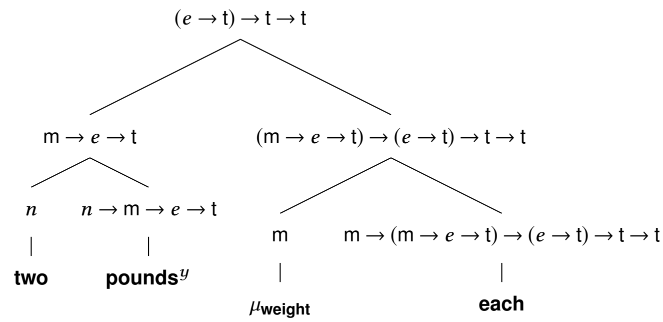
(43) a. $\textbf{each}_{x,y} := \lambda m \lambda f \lambda c.c(f\ m) \,\overline{\wedge}\, \mathsf{dm}_{x,y}\ m$ $\quad$ $\mathsf{m} \to (\mathsf{m} \to \alpha) \to ((\alpha \to \mathsf{t}) \to \mathsf{t})$

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

As already can be seen in (43-a) and (43-b), after turning a GQ (or predicate) into a higher-order



(a) a higher order dynamic GQ



(b) a higher order measure phrase predicate

Figure 3.8: Binominal *each* gives rise to a higher-order meaning

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 (43-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 (44).

(44) $\textbf{two many}^{y}\ \textbf{movies}\ \mu_{\textbf{card}}\ \textbf{each}_{x,y} =$

$$\lambda c.c\Big(\lambda P.\exists y \wedge \mathsf{movie}\ y \wedge \mu_{\mathsf{card}}\ y = \langle 2, \mu_{\mathsf{card}}, y\rangle \wedge P\ y\Big) \mathbin{\overline{\wedge}} \mathsf{dm}_{x,y}(\mu_{\mathsf{card}})$$

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

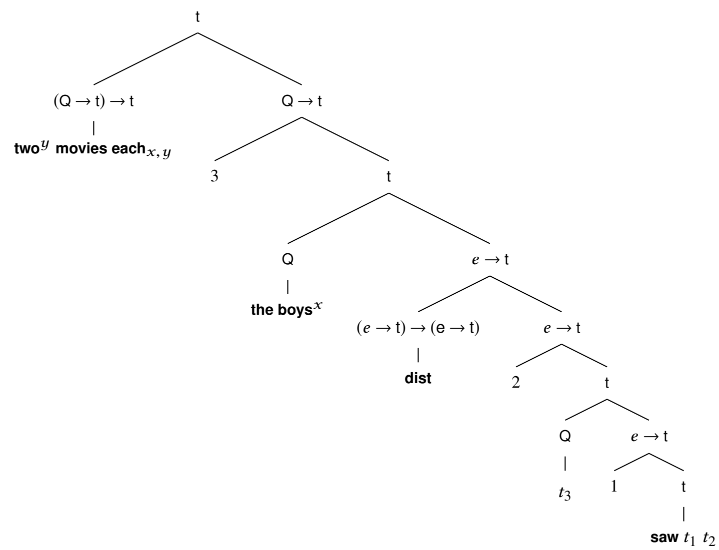


Figure 3.9: Scope taking of a higher order dynamic GQ

outside the scope of this function. Figure 3.9 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 3.3 for the definite NP *the boys*, the verb *saw* and the covert distributivity operator, we obtain the final meaning of the LF, as shown in (45).

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

Table 3.3: Definite NPs, verbs and the distributive operator

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

a. $\beta = \lambda Q.\ \mathsf{max}^{x}\left(\mathsf{boys}\ x\right) \wedge \delta_{x}\left(\mathsf{atom}\ x\ \wedge\ Q\left(\lambda u.\ \mathsf{saw}\ u\ x\right)\right)$

b. $\alpha = \mathsf{max}^{x}\left(\mathsf{boys}\ x\right) \wedge \delta_{x}\left(\begin{array}{l}\mathsf{atom}\ x\ \wedge\ \exists y\ \wedge\ \mathsf{movie}\ y\ \wedge \\ \mu_{\mathsf{card}}\ y = \langle 2, \mu_{\mathsf{card}}, y\rangle\ \wedge\ \mathsf{saw}\ y\ x\end{array}\right) \overline{\wedge}\ \mathsf{dm}_{x,y}(\mu_{\mathsf{card}})$

As shown in (45), the split scope mechanism allows *two movies* to scope inside the distributivity operator but dm to scope outside the distributivity operator. The 'association-with-distributivity' effect is clearly seen in the dm test in (45-b). The test bears an index x, which is the same index borne by the distributivity operator, i.e., the variable that stores values based on which an info-state is split up into sub-states to check for distributivity.

To test for dm, we first assemble the distributivity update. Assuming a scenario in which three boys each saw two different movies, the output of the distributivity update can be visualized in Figure 3.10.

The monotonic measurement constraint, spelled out in (46), 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, 7 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.

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

a. $H = G$ and

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

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

For concreteness, let's consider two info-states, shown in Figure 3.11, that verify dm. In info-state

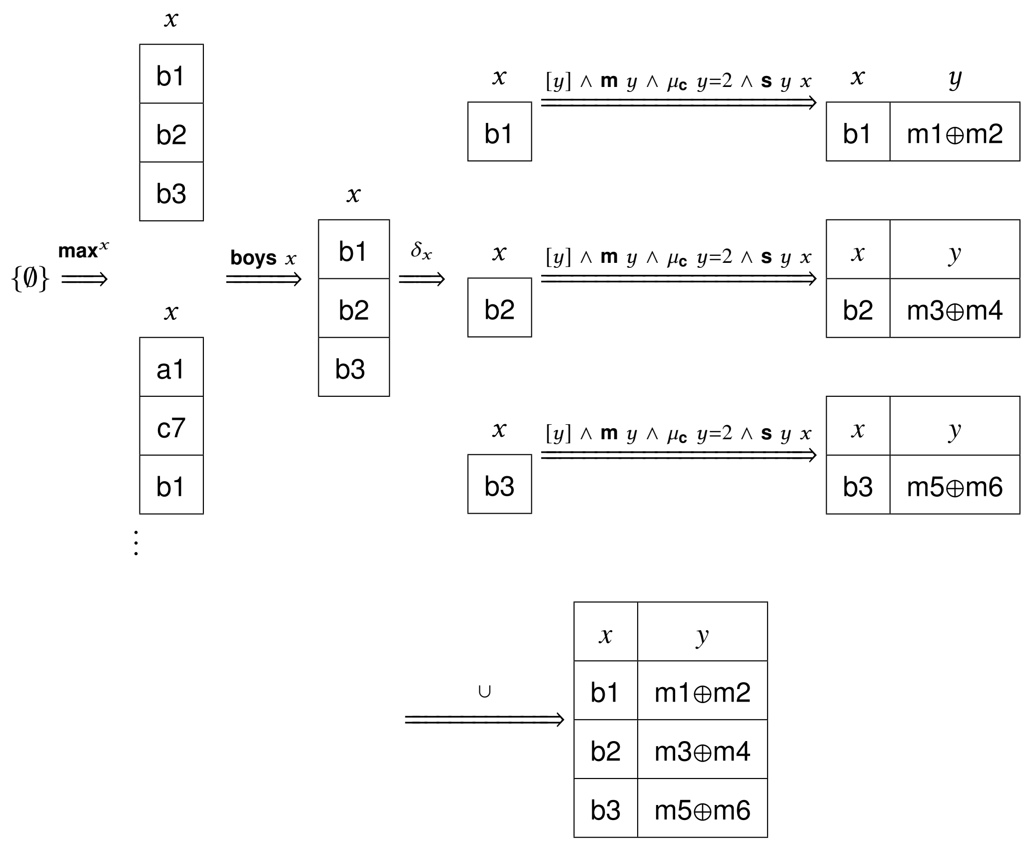


Figure 3.10: Distributivity update

G, three boys 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 boys seeing two identical movies but a third boy 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 does not vary across the distributivity dependency, as in Info-State G''. in Figure 3.12, 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*.

G	x	y
g_1	b1	m1 $\oplus$ m2
g_2	b2	m3 $\oplus$ m4
g_3	b3	m5 $\oplus$ m6

Info-state G

$\mu_{\textbf{card}}(\bigoplus A|_{x \in \{\text{b1}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m12}\rangle$

$\mu_{\textbf{card}}(\bigoplus G|_{x \in \{\text{b2}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m34}\rangle$

$\mu_{\textbf{card}}(\bigoplus G|_{x \in \{\text{b3}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m56}\rangle$

$\mu_{\textbf{card}}(\bigoplus G|_{x \in \{\text{b1, b2}\}}\ y) = \langle 4, \mu_{\textbf{card}}, \text{m1234}\rangle$

$\mu_{\textbf{card}}(\bigoplus G|_{x \in \{\text{b1, b3}\}}\ y) = \langle 4, \mu_{\textbf{card}}, \text{m1256}\rangle$

$\mu_{\textbf{card}}(\bigoplus G|_{x \in \{\text{b2, b3}\}}\ y) = \langle 4, \mu_{\textbf{card}}, \text{m3456}\rangle$

$\mu_{\textbf{card}}(\bigoplus G|_{x \in \{\text{b1, b2, b3}\}}\ y) = \langle 6, \mu_{\textbf{card}}, \text{m123456}\rangle$

G'	x	y
g'_1	b1	m1 $\oplus$ m2
g'_2	b2	m1 $\oplus$ m2
g'_3	b3	m3 $\oplus$ m4

Info-state G'

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b1}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m12}\rangle$

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b2}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m12}\rangle$

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b3}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m34}\rangle$

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b1, b2}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m12}$

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b1, b3}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m1234}\rangle$

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b2, b3}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m1234}\rangle$

$\mu_{\textbf{card}}(\bigoplus G'|_{x \in \{\text{b1, b2, b3}\}}\ y) = \langle 2, \mu_{\textbf{card}}, \text{m1234}\rangle$

Figure 3.11: Info-states that verify *the boys saw two movies each*

When the measure phrase is a predicate, as in (47-a) and (47-b), the measure phrase does 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 (47-a) and the relevant coffees for (47-b)).

(47) a. The angles are 60 degrees each.

b. *The coffees are 60 degrees each.

The corresponding monotonic measurement constraints have a similar form, as shown in (48), differing only respect to whether the values stored in x are angles or coffees, and whether the measure function measures angle degree or temperature.

(48) $G[\![\mathsf{dm}_x(\mu_{\text{angle/temp}})]\!]H = \mathbb{T}$ iff

a. $H = G$ and

b. $\forall A, A' \subseteq G\,x.\ A \subseteq A' \rightarrow \mu_{\textbf{angle/temp}}\left(\bigoplus G|_{x \in A}\ x\right) \leq \mu_{\textbf{angle/temp}}\left(\bigoplus G|_{x \in A'}\ x\right)$

c. $\exists B, B' \subseteq G\,x.\ \mu_{\textbf{angle/temp}}\left(\bigoplus G|_{x \in B}\ x\right) \neq \mu_{\textbf{angle/temp}}\left(\bigoplus G|_{x \in B'}\ x\right)$

G''	x	y
g_1	b1	m1 ⊕ m2
g_2	b2	m1 ⊕ m2
g_3	b3	m1 ⊕ m2

Info-State G''

$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b1}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$
$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b2}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$
$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b3}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$
$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b1, b2}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$
$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b1, b3}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$
$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b2, b3}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$
$$\mu_{\textsf{card}}(\bigoplus G''|_{x\in\{\textsf{b1, b2, b3}\}}\ y) = \langle 2, \mu_{\textsf{card}}, \textsf{m12}\rangle$$

Figure 3.12: An info-state that fails to verify *the boys saw two movies each*

As shown in the info-states in Figure 3.13, it is possible to satisfy dm if the measure function is extensive, as in the case of $\mu_{\textsf{angle}}$ (Info-State H), but not if the measure function is non-extensive, as in the case of $\mu_{\textsf{temp}}$ (Info-State H').[13]

3.4.4 Interim summary

I have demonstrated how to translate dm as an output constraint in DPlLM, a dynamic plural logic enriched with domain pluralities and measure functions but otherwise faithful to van den Berg (1996) (with the exception of negation, see Section 2.2.2) of Chapter 2. 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 constraint makes crucial use of the internal structure of this dependency.

In the next few sections, I discuss two extensions of the current study. The first extension

[13]When the measure function is intensive, there is no way to satisfy dm. However, when the measure function is extensive, whether or not dm is satisfied is context-dependent, as it matters what values are associated with the variable being measured.

H	x
g_1	a1
g_2	a2
g_3	a3

Info-State H

H'	x
g_1	c1
g_2	c2
g_3	c3

Info-State H'

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a1}\}}\ x) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a1}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c1}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c1}\rangle$$

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a2}\}}\ y) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a2}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c2}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c2}\rangle$$

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a3}\}}\ y) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a3}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c3}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c3}\rangle$$

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a1, a2}\}}\ y) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a12}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c1, c2}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c12}\rangle$$

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a1, a3}\}}\ y) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a13}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c1, c3}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c13}\rangle$$

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a2, a3}\}}\ y) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a13}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c2, c3}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c23}\rangle$$

$$\mu_{\textsf{angle}}(\bigoplus H|_{x \in \{\textsf{a1, a2, a3}\}}\ y) = \langle 60^\circ, \mu_{\textsf{angle}}, \textsf{a123}\rangle \qquad \mu_{\textsf{temp}}(\bigoplus H'|_{x \in \{\textsf{c1, c2, c3}\}}\ x) = \langle 60^\circ F, \mu_{\textsf{temp}}, \textsf{c123}\rangle$$

Figure 3.13: Info-states illustrating Extensive Measurement Constraint

takes up the interaction between binominal *each* and negation, with the goal of showing that their interaction follows from the dynamic framework we are using. The second extension generalizes the monotonicity constraint to cover the event differentiation condition of adverbial and determiner *each*, pointed out in studies such as Vendler (1962), Beghelli and Stowell (1997), Tunstall (1998) and Brasoveanu and Dotlačil (2015).

Following the two extensions, I offer a comparison of the proposal developed in this study and proposals developed in previous studies.

3.5 Extension 1: Negation

Negation in dynamic semantics has interesting properties. In DRT/FCS/DPL, negation is a static closure operator, which does not allow variables introduced in its scope to support anaphors outside its scope (Heim 1982, Groenendijk and Stokhof 1991, Kamp and Reyle 1993). Translating its definition into DPlLM gives rise to (49).

(49) **Static negation**

$$G[\![\neg\phi]\!]H = \mathbb{T} \ \text{ iff } G = H \ \ \& \ \ \neg\exists K : G[\![\phi]\!]K$$

Importantly, if an indefinite occurs in the scope of negation, its dynamic effect is not accessible outside the scope of negation. For this reason, cross-sentential anaphora is predicted to be impossible, as shown in (50).

(50) a. John does not own a car. #It's red.

b. Nobody talked to a man. #He left.

Given the well-documented behavior of negation in dynamic semantics, if binominal *each* indeed makes reference to (the structure of) a distributivity dependency via dynamic binding, as proposed in this study and in recent studies such as Champollion (2015) and Kuhn (2017), one wonders if it interacts with negation as predicted by dynamic semantics and the treatment of negation. This turns out to be a slightly more involved question, given the split scope behavior of a noun phrase hosting binominal *each* and the fact that the monotonic measurement constraint is defined as an output constraint. However, after unpacking the complexities, I demonstrate that binominal *each* indeed interacts with negation as predicted by the definition of negation as a static closure operator.

In a simple sentence with binominal *each* and negation like (51), there are four scopal positions for negation, as indicated in Figure 3.14.

(51) The students didn't see one movie each.

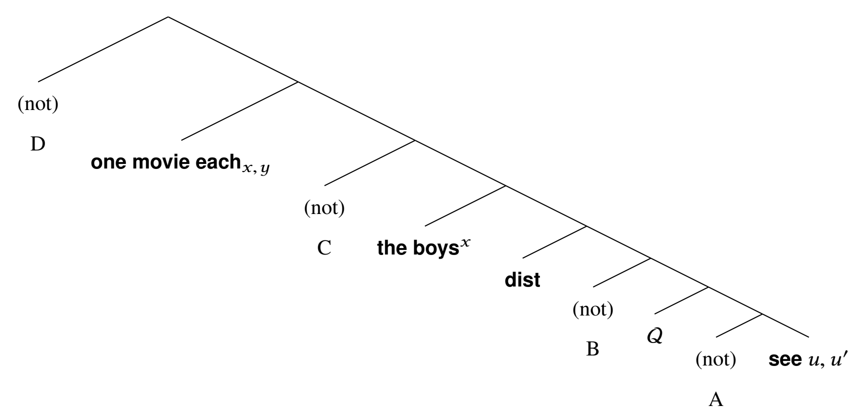


Figure 3.14: Four scopal possibilities

According to the definition of static negation, there is no well-formed interpretation if negation takes scope anywhere between the higher-order dynamic GQ and the GQ trace. In other words, B and C are not possible scope positions for negation. In position B, the dynamic effect stemming from the GQ trace (more precisely, the reconstructed GQ to the trace position) is blocked outside the scope of negation, as shown in (52); in position C, the dynamic effect stemming from both the reconstructed GQ trace and the distributivity key is blocked, as shown in (53).

(52) B. $\mathsf{max}^x\big(\mathsf{boy}\ x\big) \wedge \delta_x\left(\mathsf{atom}\ x\ \wedge\ \neg\left(\begin{array}{l}[y]\ \wedge\ \mathsf{movie}\ y\ \wedge \\ \mu_{\mathsf{card}}\ y = 1\ \wedge\ \mathsf{saw}\ y\ x\end{array}\right)\right) \overline{\wedge}\ \mathsf{dm}_{x,y}(\mu_{\mathsf{card}})$

‘For each boy, it is not true that he saw any movie. The measurement of movies is positively correlated with the number of boys.’

(53) C. $\neg\left(\mathsf{max}^x\big(\mathsf{boy}\ x\big) \wedge \delta_x\left(\mathbf{atom}\ x\ \wedge\ \begin{array}{l}[y]\ \wedge\ \mathsf{movie}\ y\ \wedge \\ \mu_{\mathsf{card}}\ y = 1\ \wedge\ \mathsf{saw}\ y\ x\end{array}\right)\right) \overline{\wedge}\ \mathbf{dm}_{x,y}(\mu_{\mathsf{card}})$

‘Not every boy saw a movie. The measurement of movies is positively correlated with the number of boys.’

How do we know that (52) and (53) are indeed out? Ideally, we should detect plain unacceptability. However, due to the availability of scopal option D, which results in a weaker reading, we only observe the lack of the **dm** inference instead of plain unacceptability in these cases. This is because a scenario that verifies (52) or (53) (without **dm**) also verifies the reading generated from having negation in position D. Moreover, **dm** can be negated or sometimes be ignored when negation is in position D. Therefore, our best evidence that (52) and (53) are indeed out comes from the fact that in situations where no boy read any book (B), or not every boy read a book (C), there is no pressure for **dm** to hold. In other words, we simply judge the sentence to be true in these situations regardless of the status of **dm**. This is suggestive that **dm** cannot be evaluated in these scope configurations.

Returning to the remaining scopal options A and D. Interpreting negation in positions A and D gives rise to well-formed interpretations. We begin with A, the simpler case. If negation takes the narrowest scope, it does not interfere with the dynamic effects associated with the distributivity key or the dynamic GQ trace, as illustrated in (54–i). As a result, **dm** can be successfully tested. This can be seen in the paraphrased in plain English in (54–ii), as well as in the sample follow-up utterance in (54–iii).

(54) A. (i) $$\textbf{max}^x\big(\textbf{boy}\ x\big) \wedge \delta_x \left(\begin{array}{l} \textbf{atom}\ x\ \wedge\ [y]\ \wedge\ \textbf{movie}\ y\ \wedge \\ \mu_{\textbf{card}}\ y = 1\ \wedge\ \neg\textbf{saw}\ y\ x \end{array} \right) \overline{\wedge}\, \textbf{dm}_{x,y}(\mu_{\textbf{card}})$$

(ii) For each boy, there is a movie that he failed to see. By the way, the measurement of the movies is positively correlated with the number of boys.

(iii) Mary didn't see *Avatar*, John didn't see *Matrix*, and Susan didn't see *Kungfu*.

If negation is to take widest scope over both the asserted distributivity and the output constraint **dm**, the resulting interpretation may be false or undefined, depending the particular truth values of the assertion and the constraint.

(55) D. (i) $$\neg\left(\textbf{max}^x\big(\textbf{boy}\ x\big) \wedge \delta_x \left(\begin{array}{l} \textbf{atom}\ x\ \wedge\ [y]\ \wedge\ \textbf{movie}\ y\ \wedge \\ \mu_{\textbf{card}}\ y = 1\ \wedge\ \textbf{saw}\ y\ x \end{array} \right) \overline{\wedge}\, \textbf{dm}_{x,y}(\mu_{\textbf{card}})\right)$$

(ii) It's not true (that every boy saw two movies and by the way, the measurement of movies is positively correlated with the number of boys).

Using ϕ to represent the asserted content, i.e., the distributivity update, ψ to represent the output constraint, i.e., **dm**, $\overline{\wedge}$ to represent the outcome of the ϕ as constrained by ψ, and $\neg$ to represent the predicted outcome of negation over the complex meaning, Table 3.4 summarizes the possible interpretation of $\neg$.

ϕ	ψ	$\overline{\wedge}$	$\neg$
1	1	1	0
1	0	#	#
0	#	#	#

Table 3.4: Negation over complex meaning

Let us begin with the first two rows. When both the distributivity evaluation and the monotonic measurement constraint evaluation are 'true', negation is evaluated to 'false', in accordance with native speakers' intuition that the sentence in (56) is simply a false statement.

(56) *In a scenario in which every boy watched a different movie:*

The boys didn't watch one movie each.

When distributivity is evaluated to 'true' but the monotonic measurement constraint is evaluated to 'false', negation is evaluated to 'undefined'. The cases in (57) and (58) support this prediction.

(57) *In a scenario in which all the boys watched one and the same movie:*

#The boys didn't watch one movie each.

(58) *In a scenario in which all the cocktails are exactly 60 degrees Fahrenheit.*

#The cocktails aren't 60 degrees (Fahrenheit) each.

When distributivity is evaluated to 'false', the output is an empty set, so the monotonic measurement constraint cannot be tested. As a result, $\overline{\wedge}$ is evaluated to 'undefined'. Applying negation to an undefined output is also predicted to be 'undefined'. This is where native speakers' intuition differs from the prediction. Intuitively, if distributivity is false, speakers will simply evaluate the negated sentence to 'true', rather than 'undefined'. The following example demonstrates the judgment.

(59) *In a scenario in which not every boy watched a movie:*

The boys didn't watch one movie each.

This discrepancy between what the semantics predicts and what native speakers perceive in this particular type of examples is actually a more general phenomenon in dynamic semantics.[14] Consider (60) first. Imagine a situation in which no man came in. The evaluation of *a man came in* is false in this situation. The result is the absurd state, i.e., having no info-state in the output. Consequently, he_x in the subsequent clause has no way to refer to a^x *man* as its antecedent, leading to an undefined evaluation. When the outer negation is evaluated, it should then be evaluated to 'undefined'. However, this is not what native speakers do. In fact, they readily judge the negated sentence to be 'true', instead of 'undefined'. It seems as if negation has the ability to ignore dynamic binding failure that happens in its scope.[15]

[14] I thank Simon Charlow (p.c.) for helpful discussions on negation in dynamic semantics.

[15] In intensional versions of dynamic semantics, such as Heim 1982 and Brasoveanu 2010, the predictions are subtler than this.

(60) It's not the case that a^x man came in and he$_x$ sat down.

However, negation cannot ignore just any dynamic binding failure. For example, it cannot ignore dynamic binding failure that fails a gender presupposition, as evidence by the fact that (61) is evaluated to 'undefined'.

(61) #It's not the case that a^x man came in and she$_x$ sat down.

Returning to the last row of Table 3.4. Native speakers evaluate the negated sentence with binominal *each* to be 'true', rather than the predicted 'undefined', precisely for the same reason. That is, negation has the ability to ignore dynamic binding failure in its scope. Similarly, negation only has the ability to ignore dynamic binding failure in this case, but not other anomalies. For example, if the measurement involved in the host becomes non-extensive, undefinedness resurfaces, as demonstrated in (62).

(62) a. *The boys didn't walked 3 miles per hour each.
b. *The drinks are not 60 degrees each.

3.6 Comparisons with and connections to previous studies

3.6.1 Studies in the PCDRT framework

Henderson (2014) investigates distributive numerals in Kaqchikel and conclude that they should be analyzed as imposing a post-suppositional plurality condition (known as evaluation-level plurality) on the functional dependency arising from distributive quantification, which he modeled using info-states from PCDRT. Champollion (2015) later extends Henderson's analysis to binominal *each* (with a novel river metaphor to help illustrate PCDRT). Kuhn (2017) modifies Henderson's analysis, replacing post-suppositions with a scope-taking mechanism, and allowing distributive numerals (noun phrases with binominal *each* included) to induce distributivity. However, the core of Henderson's analysis, namely, that distributive numerals contribute an evaluation-level plurality condition, is shared in both Champollion (2015) and Kuhn (2017).

Since the evaluation-level plurality condition is a major point of departure between the present study and previous studies in the PCDRT tradition, let me introduce it with a concrete sentence like (63). This sentence can be translated into DPlLM as in (64). The only bit that cannot be interpreted in DPlLM is the evaluation-level plurality condition in the last conjunct. Let me define it in (65), using Henderson's definition.

(63) The students hugged one dog each.

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

Definition Evaluation-level plurality (PCDRT)

(65) $G[\![y > 1]\!]H = \mathbb{T}$ iff $G = H$ and $|\{h\ y \mid h \in H\}| > 1$

h in (65) 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 at the value level. To see this, consider the definition of value dependence, repeated below from Definition 8 of Chapter 2.

Definition 8 (Value dependence)

y is value-dependent on x in an information state G iff there are $a, b \in G\ x . G|_{x=a}\ y \neq G|_{x=b}\ y$

If we set the variable x to store all the relevant values in the distributivity key and y to stores the values introduced by a distributive numeral. 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 the same requirements. 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 distributivity key is not used explicitly. However, it is used implicitly, as

Value dependence is useful for modeling the variation requirement, but does not capture the measurement-sensitivity of binominal *each*, which subsumes the counting quantifier requirement and the extensive measurement requirement. The monotonic measurement constraint, by contrast, captures both the variation requirement and the measurement sensitivity.

It can be recast in terms of structural dependence but it does not need to be. The

However, it cannot handle requirements that track the internal structure of a functional dependency, such as dependence between individuals and measurements of individuals, i.e., degrees, as discussed in the previous chapter. As has been established in Section 3.2, binominal *each* makes crucial reference to measurement. For this reason, value dependence via evaluation-level cardinality cannot be extended to handle the measurement-sensitive nature of binominal *each*.

In addition to the primary difference between value dependence and structural dependence, 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 Champollion (2015), evaluation-level plurality it is analyzed as a delayed at-issue test. 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 empirical 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’ to the constraint formula, which is precisely the output of distributive quantification. I call the constraint an output constraint because I intend for it to constrain the output of distributive quantification.

In addition, this study has adopted a higher order meaning approach to model delayed evaluation, following Charlow (to appear). Henderson (2014) follows Brasoveanu (2013) and uses a postsupposition instead and the assumption carries over to Champollion (2015). Lastly, Kuhn (2017) uses ordinary scope-taking without higher order meaning to model delayed evaluation. The merits and shortcomings of these strategies are discussed in Section 3.4 (see also Charlow, to appear).

A third difference lies in the dynamic logic in which the constraint giving rise to the variation requirement is couched. Studies in the PCDRT tradition make use of PCDRT, a dynamic plural logic with domain pluralities, dependence-introducing variable introduction, and distributive evaluation of lexical relations. The present study is couched in DPlLM, a dynamic plural logic with domain pluralities and a collective evaluation of lexical relations, but crucially no dependency-free variable introduction. The two logics make distinct predictions regarding whether dependencies could in principle occur without distributivity. In particular, PCDRT allows it while the present framework does not. As a consequence, PCDRT-theoretic studies predict that the variation component may exist independently of distributivity, while the present framework predicts a close connection between distributivity and the variation requirement.[16]

I must add that the monotonic measurement constraint, the constraint designed to replace evaluation-level plurality, can be reformulated with minor changes to adapt to a PCDRT-style dynamic plural logic just as easily. Ultimately, whether a PCDRT-style logic or a DPlLM-style logic should be chosen to couch the monotonicity constraint should be based on empirical considerations. In particular, if the dependency introduced by random assignment turns out to be very useful, as argued in Brasoveanu (2007, 2008), then a PCDRT-style logic should be favored. However, there is at least some initial evidence, reported in Champollion et al. (forthcoming), that the original empirical motivations for the dependency-introducing random assignment considered in Brasoveanu (2007, 2008) may be accounted for without the machinery used in PCDRT.

3.6.2 Studies in static semantics

There is a vast literature on binominal *each* couched in static semantics. It is beyond the scope of the present chapter 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

[16] Henderson (2014) is aware that dependence-introducing variable introduction has the potential to over-generate the licensors for distributive numerals. To avoid over-generation, he requires variables to have only one value at the evaluation level (i.e., $x = 1$, or to be singular 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 in Section 2.2.3 of Chapter 2, which is based on van den Berg (1996), can be seen as a more 'automatic' way of getting rid of undesirable dependencies—they are not generated in the first place.

chapter.

An early study on binominal *each* is Link (1987). He set the stage for treating binominal *each* as a distributivity operator, which 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 chapter, 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 distributive numerals in Tlingit to binominal *each*, arguing that *each* is both a distributivity 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 extensive 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.[17]

Many studies also share the intuition that *each* is a marker of quantificational dependence or subordination. Choe (1987a) 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 subordination. However, in the present study, the core contribution is a variation component formalized in terms of a monotonic

[17]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.

measurement constraint. A separate constraint is needed because quantificational subordination is a *necessary but not sufficient* condition for using binominal *each*.

As a final note, I would like to relate the present study 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 the framework of Conceptual Semantics, which differs from the framework used in this work quite substantially. However, the core of the idea of structure-preserving binding resonates with the notion of the monotonicity constraints developed here.

3.6.3 Other ways to model functional dependencies

Besides using sets of assignments, a few other approaches have been developed to model functional dependencies arising from distributive quantification. I discuss two options in this section.

The first approach is to use **Skolemized choice functions**. Some notable studies arguing for a Skolem function treatment of universal quantification is Huang (1996) and Solomon (2011). According to these studies, universal quantification (closely related to distributive quantification) has the effect of functionalizing an expression in its scope. In Solomon (2011), the expression is an indefinite, while in Huang (1996), it may be an indefinite or an event. Abstracting away from the compositional details, what these scholars suggest is essentially representing a sentence like (66-a) as (66-b), where f is a Skolem function mapping each girl to a book she read.

(66) a. The girls each read a book.

b. $\exists f \forall x. x \leq \bigoplus {}^*\mathsf{girl} \,\&\, \mathsf{atom}(x) \rightarrow \mathsf{book}(fx) \,\&\, \mathsf{read}(fx)(x)$

The monotonicity constraint imposed by binominal *each* can be defined using a Skolem function:

(67) $\mathsf{dm}(f) := \forall A, A' \subseteq \mathsf{Dom}(f). A \subseteq A' \rightarrow \mu \bigoplus \{fx \mid x \in A\} \leqslant \mu \bigoplus \{fx' \mid x' \in A'\}$

The constraint in (67) can be directly conjoined with the contribution of distributive quantification in (66-b). f will be correctly identified to be the Skolem function contributed by the distributive quantification.

Despite its initial success, however, this approach faces some challenges. First, when the indefinite in the share happens to be a downward monotone quantifier, f may happen to be a defective function. For example, (68-a) is grammatical and compatible with a situation in which some of the girls did not read any book. Its corresponding interpretation involving a Skolem function, as formulated in (68-b), is predicted to be defective in a situation with some girls not reading any book. This is because the Skolem function f must choose some book from the set, predicting that every girl must read some book. In addition, f is not appropriately maximized, so (68-b) is true even in a situation in which the girls each read more than one book, as one can always find a (less informative) function that assigns no more than one book to each girl.

(68) a. The girls read no more than one book each.

b. $\exists f \forall x. x \leq \bigoplus {}^{*}\mathsf{girl}\ \&\ \mathsf{atom}(x) \rightarrow \mathsf{book}(fx)\ \&\ |(fx)| \leqslant 1\ \&\ \mathsf{read}(fx)(x)$

The two problems identified above can be resolved by letting the Skolem function take narrow scope relative to distributivity and a negation operator contributed by the downward monotone quantifier.

(69) $\forall x. x \leq \bigoplus {}^{*}\mathsf{girl}\ \&\ \mathsf{atom}(x) \rightarrow \neg\exists f.\mathsf{book}(fx)\ \&\ |(fx)| > 1\ \&\ \mathsf{read}(fx)(x)$

However, doing so renders it impossible for the monotonicity constraint to access f in (69), as it is embedded under negation. For the constraint to access f, it has to occur inside the scope of negation. However, the resulting interpretation would become: for every girl, there is no function that (i) maps the girl to more than a book she read and (ii) overall the girls read a variety of books. The is true when each girl read the same book, obviously not what (68-a) means.

Another problem with the Skolem function approach concerns the difficulty in assembling a Skolem function compositionally. Huang (1996) has to make special assumptions to fix the domain and range of the function as they are being pulled from distant parts of a sentence—the domain is pulled from the distributivity key, and the range from the indefinite (or whatever expression that functionally depends on the distributivity key).[18] Solomon (2011) offers a more compositional

[18]For example, Huang (1996) has to rely on two assumptions to set the domain and range of the function: (i) an indefinite in the distributed share contributes only a variable and not existential quantification, à la DRT, and (ii) *every/each* binds this variable.

account, at the cost of upgrading the semantics of distributive quantifiers.

If Skolem functions coming from distributive quantification is not ideal, one may wonder if attributing the source of Skolem functions to indefinites works any better. After all, choice functions have been widely used to analyze indefinites taking exceptional scope (Hintikka 1973, Kratzer 1998, Matthewson 1999, Winter 2002, Schwarz 2001). Their Skolemized versions, known sometimes as Skolemized choice functions, can model indefinites interpreted in the scope of another quantifier, either because it contains a pronoun bound by the the latter or because it is assumed to be functionally dependent on the latter. Relatedly, Milačić et al. (2015) have argued that Skolem functions be used for modeling the semantics of distributive numerals.

However, this proposal does not line up with empirical facts. Indefinites that have been widely accepted as having a functional interpretation, such as *a certain NP*, do not support binominal *each*, as shown in (70). By contrast, noun phrases that do not have exceptional scope properties and hence are not typically analyzed as functional indefinites, such as counting quantifiers with modified numerals (cf. Abels and Martí 2010), as shown in (71), do support binominal *each*.

(70) *The girls read a certain book each.

(71) The girls read at least/most two books each.

Therefore, we can conclude that Skolem functions contributed by indefinites is neither sufficient (70) nor necessary (71) for supporting binominal *each*.

Event semantics

Alternatively, studies have tried to encode the functional dependency of distributivity using Event Semantics, as done in Schein (1993), Lasersohn (1995), Landman (2000), and Champollion (2010, 2017). One of the reasons Event Semantics is chosen for this job is because events have (mereological) parts, which can, in principle, be used to model the structure needed for building the functional dependency of distributivity. As an illustration, I use Schein's semantics for distributive quantifiers shown in (72).

(72) Schein's semantics for a distributive quantifier

every girl $:= \lambda P.\lambda e.\forall x[\textsf{girl}(x)](\exists e'[e' \leq e](P(x)(e'))])$

When this quantifier is used in a sentence like (73-a), it gives rise to the interpretation in (73-b).

(73) a. Every girl left.

b. $\exists e(\forall x[\textsf{girl}(x)](\exists e'[e' \leq e](\textsf{leave}(x)(e'))))$

Here, e is a sum event with parts. At least some of these parts are events in which a girl left.[19] For each girl, her leaving event takes up a part of e. Because of the availability of the sum event e, one can in principle define ways to extract the girl-event dependency in (73-b).

I do not have a strong objection to using Event Semantics for modeling distributivity. As long as one can successfully retrieve the relevant participants that form a functional dependency from events, the main ideas developed in this dissertation can be largely transported in an event-based framework.

3.7 Conclusion

In this work, I have borrowed the insight from previous studies that distributivity makes available a set of 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). Following many recent studies, this dependency is modeled with help of a dynamic plural logic. The particular version used in chapter is a hybrid of van den Berg (1996)'s DPlL and Brasoveanu (2008)'s PCDRT. More details about this logic are given in the previous chapter.

In addition, I have argued that binominal *each* piggybacks on this dependency, and introduces a monotonicity constraint requiring that the measurement of the values associated with its host tracks the part-whole relation of the dependency. I have demonstrated how the monotonicity constraint shed light on three generalizations on binominal *each*: the variation requirement, the counting quantifier requirement and the extensive measurement requirement.

[19] Schein's semantics is too weak, as pointed out in Ferreira (2005) and Champollion (2010). This is because there is no restriction placed on e other than that it consists of a set of subevents each of which is a girl-leaving event. Besides these subevents, it may contain many other events that are not relevant. Both Ferreira (2005) and Champollion (2010) have provided ways to improve Schein's event-based distributivity.

Moreover, I have also shown that the monotonicity constraint can be generalized to account for the event differentiation condition associated with adverbial and determiner *each*. The generalization helps us see that even ordinary distributivity markers, such as determiner and adverbial *each*, may encode constraints on distributivity.

Lastly, I have demonstrated that a dynamic treatment of binominal *each* makes correct predictions about its interactions with negation, justifying the use of dynamic semantics in this area of research.

4

CANTONESE *saai*

4.1 Introduction

Previous studies have identified distributivity markers that require *obligatory co-variation* of expressions in the distributivity share relative to the distributivity key. Well known examples of this type of distributivity markers include: markers of distributive numerals/indefinites, which are found in Georgian, Hungarian, Telegu, Kaqchikel, ASL, and many other languages (Farkas 1997, Balusu 2005, Henderson 2014, Kuhn 2017, a.o.), English binominal *each* (Choe 1987b, Safir and Stowell 1988, Champollion 2015), and Mandarin *ge* (Lin 2004, Lee et al. 2009a, Li and Law 2016).

One may wonder if there is any distributivity marker that does just the opposite, namely, requiring expressions in the distributivity share to *not co-vary* with the distributivity key. In this study, I show that such a distributivity marker can be found in Cantonese.

The distributivity marker that exhibits this property is the post-verbal distributivity suffix *saai*, which has been previously taken up in Lee (1994), Tang (1996) and Lee (2012). These studies have already observed that *saai* strongly resists indefinites showing up in the object position of a transitive verb suffixed by *saai*. Lee (1994) and Tang (1996) even suggested building in a definiteness or specificity requirement in the semantics of *saai* to explain this incompatibility.

Against this background, the contribution of this study is three-fold.

- At the empirical front, I show that *saai*'s resistance of indefinites is interpretation-based—as long as an indefinite does not co-vary with the distributivity key, the incompatibility goes away. The resistance against co-variation is demonstrated to generalize to disjunction (section 3.2), as well as to measure phrases (section 4.5).

- At the theoretical front, I argue that definiteness and specificity are not adequate notions for accounting for the property of *saai*-distributivity (**??**). Instead, I propose that the ban on co-variation can be understood in terms of *independence*. I explored two analyses, one relying on scope (section 4.3) and the other relying on an independence constraint (section 4.4), to model independence in *saai*-distributivity. The independence constraint is further argued to hold not at the value level, i.e., requiring a lack of co-variation with the distributivity key, but at the structure level, i.e., requiring a lack of co-variation with the internal mereological structure of the distributivity key (section 4.5).

- Finally, by identifying *saai* as a distributivity marker indicating independence, I have enriched the typology of distributivity markers: there are distributivity markers signaling (structural) dependence (as in the case of English *each* and Mandarin *ge*) as well as distributivity markers signaling (structural) independence.

4.2 The distribution of *saai*

4.2.1 Establishing *saai* as a distributivity marker

Saai is a verbal suffix indicating distributivity, according to Tang (1996), Lee (2012), and Lei (2017). These authors have also provided a variety of evidence to support *saai*'s status as a distributivity marker, including *saai*'s requirement for a plural distributivity key, as well as its interactions with collective predicates and mixed predicates. I briefly review these pieces of evidence below to show that *saai* has typical properties of a distributivity marker.

According to Tang (1996), *saai* requires a distributivity key that exhibits 'divisibility', in the sense that the distributivity key must be divisible into a plurality of proper parts. This is a common characteristic of distributivity markers and can be seen as a ban against distributive quantification operating on a singleton domain.

(1) Keoidei/*keoi zau-saai.
they/he leave-SAAI
'They/*he each left.'

A immediate qualification is that the distributivity key can be in the singular form, as shown in (2). What is important is that it be divisible into proper subparts to feed distributive quantification. This property is shared by a more widely studied distributivity marker *dou* in Mandarin and has been formalized in terms of a cover-based semantics for distributivity in Lin (1998a) (see also Schwarzschild 1996).

(2) Bun-syu sap-saai.
CL-book wet-SAAI
'The whole book is wet.'

Tang (1996) further suggests the contrast in (3) and (4) to establish *saai*'s role as a distributivity marker. According to Tang (1996), *git-zo fan* 'got married' in (3) is ambiguous between a collective interpretation (3-a) and a distributive interpretation (3-b). However, suffixing the verb with *saai* instead of *zo* makes the collective interpretation unavailable, as demonstrated in (4).

(3) Keoidei git-zo fan.
they get-ASP tie-ASP marriage

a. Collective: They got married with each other.

b. Distributive: They each married someone.

(4) Keoidei git-saai fan.
they get-SAAI marriage

a. *Collective: They married each other.

b. Distributive: They each married someone.

A concern with using the contrast in (3) and (4) to diagnose distributivity is that the size of the plural entity that serves as the distributivity key makes a difference to the judgment. If the plural pronoun in (4) refers to a bigger group of individuals, then the 'collective' interpretation is more easily acceptable. This is reminiscent of *all* in English, which is compatible with a type of collective predicates called *gather*-type predicates, as shown in (5).

(5) All the students gathered.

However, Champollion (2017) argues that there are two types of collective predicates: those that allow distributivity down to non-atomic units and those that resist distributivity altogether (see also Dowty 1987, Kuhn 2014). The first type of predicate is exemplified by *gather, fit together,* and *hold hands*, while the second type of predicate is exemplified by *be numerous,* and *be a large group*. Although *all* is compatible with gather-type predicates, it is incompatible with *numerous*-type predicates, suggesting that it is indeed incompatible with a genuine collective predicate.

(6) *All the students are numerous.

Saai in Cantonese patterns like *all* in this respect. It is compatible with a *gather*-type predicate (7) and incompatible with incompatible with a *numerous*-type predicate (8).[1]

(7) Keoidei zeoi-saai hai munhau.
they gather-SAAI at entrance
'The gathered at the entrance.'

(8) Keoidei jan do-*zo*/**saai*.
they people large-ASP/SAAI
Intended: 'They became numerous.'

Another test that can be used to establish *saai*'s role as signaling distributivity involves the use of the so-called 'mixed' predicates, i.e., predicates that are ambiguous between a distributive and a collective interpretation. The predicate *toi jat-bou gongkam* 'lift a piano' in (9) is an example of a 'mixed' predicate. The two interpretations are included immediately below the sentence.

(9) Di-hoksaang toi-zo jat-bou gongkam.
CL.PL-student lift-ASP one-CL piano

a. Collective: The students lifted a piano together.

b. Distributive: The student each lifted a piano.

Replacing the aspectual suffix *zo* with *saai* brings about two changes. First, it makes the collective interpretation unavailable, leaving the distributive interpretation the only viable interpretation:

[1]Mandarin *dou*, a marker that has been argued to be a distributivity marker, is also compatible with *gather*-type predicates (Lin 1998a)

(10) %Di-hoksaang toi-saai jat-bou gongkam.

CL.PL-student lift-ASP one-CL piano

a. *Collective: The students lifted a piano together.

b. Distributive: The student each lifted a piano.

Second, the sentence itself is slightly degraded for some speakers, even for the distributive interpretation. Lee (1994), Tang (1996), Lee (2012) and Lei (2017) attribute the degradedness to the requirement that a noun phrase following *saai* has to receive a specific or definite interpretation. The specificity (or definiteness) requirement, which is at the heart of this study, is discussed in more detail in the next subsection.

In summary, with help from *numerous*-type predicates and mixed predicates, I have shown that Cantonese *saai* signals distributivity. Let me immediately clarify that signaling distributivity does not equal contributing a distributivity operator. *Saai* may merely indicate the presence of a distributivity operator.

4.2.2 Interactions with indefinites and disjunction

Saai is selective about what expression may follow it and the range of interpretations a post-*saai* expression may take. This subsection documents these selectional requirements. It incorporates observations noted in studies such as Lee (1994), Tang (1996), Lee et al. (2009a) and Lei (2017), as well as observations stemming from my own research.

It is widely noted that *saai* favors definite expressions as post-*saai* objects over indefinite expressions (e.g., Lee 1994, Tang 1996, Lee 2012). Cantonese has a few ways to form definite expressions. For example, they can be formed with a demonstrative determiner followed by a classifier and a common noun, as shown in (11), as well as with a bare classifier (i.e., without a numeral) followed by a common noun, as shown in (12) (Au Yeung 1998, Cheng and Sybesma 1999, Jiang 2012). Both expressions can follow *saai* without any issue.

(11) Di-hoksaang gin-saai lei-go lousi.
CL.PL-student see-SAAI this-CL teacher
'The students each saw this teacher.' Demonstrative definite

(12) Di-hoksaang gin-saai go lousi.
CL.PL-student see-SAAI CL teacher
'The students each saw the teacher.' Classifier definite

Indefinites in Cantonese are typically introduced by numeral classifier constructions of the form [Num + Cl + N]. Numeral classifier constructions are closest to counting quantifiers in English. I call them cardinal indefinites to emphasize their indefinite nature. As shown in (13), *saai* is marked when followed by such a cardinal indefinite.

(13) %Di-hoksaang gin-saai jat-go lousi.
CL.PL-student see-SAAI one-CL teacher
'The students all saw a teacher.' Cardinal indefinite

The judgment is similar for cardinal indefinites involving a higher numeral, as shown in (14). However, sentences involving two plural arguments and *saai* have an extra layer of complexity, as they are ambiguous between a subject-distributivity reading and an object-distributivity reading. In the subject-distributivity reading, the subject is the distributivity key, whereas in the object-distributivity reading, the object is the distributivity key. This flexibility is a well-documented property of distributivity with *saai* (Lee 2012) and is discussed as a remaining issue in Section 4.7.[2]

(14) %Di-hoksaang gin-saai saam-go lousi.
CL.PL-student see-SAAI three-CL teacher
'The students all saw three teacher.'

Bare noun phrases are also sometimes treated as indefinites in Cantonese for they occur in

[2] Other quantifiers also exhibit similar markedness when the distributivity key is set to be the subject:

(i) %Di-hoksaang gin-saai daboufeng lousi.
CL.PL-student see-SAAI most teacher
'The students all saw most teachers.'

(ii) %Di-hoksaang gin-saai housiu lousi.
CL.PL-student see-SAAI few teacher
'The students all saw few teachers.'

When the distributivity key is set to be the object, the markedness disappears, because the distributed share now contains the subject and the verb. The subject is a definite expression so it is not marked.

Due to the ambiguous nature of data involving two plural arguments, I do not think they offer us the clearest clue as to what distinguishes between definite expressions and indefinite expressions in the distributed share when distributivity is marked by *saai*. For this reason, I do not delve into data involving plural indefinites and plural quantifiers.

existential sentences. However, *saai* is compatible with bare noun phrases. In other words, bare noun phrases pattern like definite expressions in not showing markedness effects when they follow *saai*.

(15) Di-hoksaang gin-saai lousi.
CL.PL-student see-SAAI teacher
'The students all saw one or more teachers.' Bare NPs

No study has addressed why bare noun phrases pattern more like definite expressions in terms of their ability to appear after *saai* (but see the analysis proposed in this study in Section 4.4.3). The contrast between (11) – (12) and (13), however, has motivated the generalization that a post-*saai* expression must be definite or specific. The literature has not formally tested whether definiteness or specificity is the relevant notion. Presumably, if the relevant notion is a definiteness requirement, then the inter-speaker variability reflects whether or not a speaker can assign a definite interpretation using an indefinite form; if the relevant notion is specificity, then the variability reflects whether or not one can assign a specific indefinite interpretation to an indefinite form.

There are a few reasons to believe that definiteness is not the right notion. First, a post-*saai* expression may enter into scope interaction with negation in ways that a definite expression cannot. For example, the indefinite in (16) may be interpreted as having wide scope relative to negation (30-a), or as having narrow scope relative to negation (30-b).[3]

(16) Di-hoksaang mou gin-saai jat-go lousi.
CL.PL-students not see-SAAI one-CL teacher
a. There is a teacher that the students did not all see.
b. There is no teacher that the students all saw.

However, a definite expression in the same position does not interact with negation. As a result, the following sentence only has one interpretation:

(17) Di-hoksaang mou gin-saai go-lousi.
CL-students not see-SAAI CL-teacher
'The students did not each see the teacher.'

[3]The narrow-scope interpretation can be further brought out by the use of *jamho* 'any' before the numeral.

Second, a post-*saai* indefinite must introduce a referent that is discourse-novel. B's answer in (18) is infelicitous because the indefinite has the same referent as the possessive noun phrase introduced in A's question.

(18) A: Gamjong-bun sun^x syu hotaai ma?
Gamjong-CL new book goodread POLQ
'Is Gamjong's new book a good read?'

B: #Hotaai aa. Di-hoksaang maai-saai jat-bun syu$_x$ lai taai.
goodread SFP CL-students buy-SAAI one-CL book to read
'Yes. The students each bought a book to read.'

However, using a definite expression in the same position is acceptable:

(19) A: Gamjong-bun sun syu hotaai ma?
Gamjong-CL new book goodread POLQ
'Is Gamjong's new book a good read?'

B: Hotaai aa. Di-hoksaang maai-saai bun-syu lai taai.
goodread SFP. CL.PL-students buy-SAAI CL-book to read
'Yes. The students all bought the book to read.'

Given the above considerations, the markedness of indefinites should not be explained in terms of a definiteness requirement on the indefinites. This leaves us with the specificity requirement. That is, the markedness of indefinites is due to the need to interpret them as specific indefinites. I provide some evidence below suggesting that specificity indeed provides a more adequate explanation. Reserving a more precise formulation of the kind of specificity involved in *saai* until Section 4.3 and 4.4, here I only use specificity in an intuitive sense: an indefinite is specific when it refers to a referent identifiable by the speaker.[4]

To begin with, it is known that indefinites with a more descriptive content have a better chance

[4]I examined all instances of the post-verbal *saai* that co-occur with a cardinal indefinite in Hong Kong Cantonese Corpus (CanCorp, Luke and Wong 2005). The findings are two-fold. First, cardinal indefinites do follow *saai* in naturally occurring discourse. Second, all the cardinal indefinites following *saai* invariably have a strong specific indefinite flavor. I provide two natural occurring data here for illustration. The data are slightly modified to reduce their length and have the dropped arguments re-introduced in parentheses to facilitate interpretation). In both sentences, the indefinite has a specific referent: a specific tank in (i) and a specific CD in (i). In other words, (i) cannot be true if the fish are in different tanks, and (i) cannot be true if the good songs are in different CDs.

(i) Dim wuzou (di-jyu) dou hai-saai jat-go gong leoibin lo.
however dirty CL-fish dou in-SAAI one-CL tank inside SFP
'However dirty, (the fish) are all inside a single tank.'[5] (ID: FC-033)

being interpreted as specific (Fodor and Sag 1982, Schwarzschild and Wilkinson 2002). Relatedly, enriching the descriptive content of an indefinite by introducing a relative clause improves its ability to co-occur with *saai*.

(20) Di-hoksaang gin-saai jat-go san lei-ge lousi.
CL.PL-student see-SAAI one-CL newly arrive-MOD teacher
'The students all saw a newly hired teachers.'

In addition, by comparing the interaction of indefinites with *saai* and the interaction of indefinites with other distributivity markers, it can be shown that indefinites following *saai* indeed receives a specific interpretation. There are two other distributivity markers in Cantonese: *cyunbou* 'all, completely' and *dou*.[6] Indefinites co-occurring with these markers do not exhibit markedness effects and may co-vary with distributive quantification, as shown below:

(21) a. Keoidei cyunbou gin-zo jat-go lousi
they all saw-ASP one-CL teacher
'They all saw a teacher (possibly different teachers). '

b. Cyunbou jan dou gin-zo jat-go lousi
all people DOU saw-ASP one-CL teacher
'All the people saw a teacher (possibly different teachers). '

However, indefinites following *saai* not only are marked for some speakers, but also may not co-vary with distributive quantification for the speakers who accept them:

(22) %Di-hoksaang gin-saai jat-go lousi.
CL.PL-student see-SAAI one-CL teacher
'The students all saw a teacher (the same teacher).'

(i) Hou go dou baai-saai lok jat-zek (dip) dou aa.
good song dou put-SAAI in one-CL CD there SFP
'The good songs are all in one (CD).' (ID: FC-109a)

[6]Cantonese *dou* is a cognate of the more famous Mandarin *dou* (Cheng 1995). *Dou* (in both languages) signals distributivity, but whether it constitutes a distributivity operator is subject to debate. While Lin (1998a) argues that *dou* is a generalized distributivity operator, many recent studies disputed this view, including Chen (2005); **?**, Xiang (2008), Liu (2016), Xiang (2016).

I am not aware of any formal analysis of Cantonese *cyunbou*, other than its partial cognate *quan* take up in Tomioka and Tsai (2005). However, the properties of *cyunbou* and *quan* are quite different. In this study, I use *all* to translate *cyunbou*. However, it must be noted that Brisson (1998, 2003) argues that *all* is not a distributivity operator but a maximality marker for removing pragmatic slack. I have not investigated *cyunbou* in enough detail to determine whether it is a genuine distributivity operator or a maximality operator akin to *all*.

Lastly, specificity can be extended to understand the interaction of *saai* and disjunction. Compare (23-a), which allows the disjunction to co-vary with the distributivity key when distributivity is marked with *cyunbou* and *dou*, with (23-b), which disallows the co-variation when *cyunbou* and *dou* are replaced by *saai*.

(23) a. Di-hoksaang cyunbou dou maai-zo Emma waatze Jane Eyre.
CL.PL-student all dou buy-ASP Emma or Jane Eyre
'The students all bought Emma or Jane Eyre.'

b. Di-hoksaang maai-saai Emma waatze Jane Eyre.
CL.PL-student buy-SAAI Emma or Jane Eyre
'The students all bought Emma or they all bought Jane Eyre.'

In summary, when distributivity is marked with *saai*, the expression following *saai* has to assume a specific interpretation. By comparing the interpretation of indefinites co-occurring with *saai* and those co-occurring with other distributivity markers, I have shown that the specificity comes from *saai* rather than from the indefinites. An important question arising from this discussion is why as a distributivity marker *saai* carries a specificity requirement. To answer this question, it is necessary to understand the type of specificity associated with *saai*, a task I take up in the next subsection.

4.2.3 Independence in distributivity

As pointed out in many studies, there is no agreed-upon definition for specificity. The main reason is because there are different types of specificity, each with its own characteristics (e.g., Farkas 2002a, von Heusinger 2002). In this section, I explore three types of specificity that may be associated with *saai*'s specificity effect. They are *epistemic specificity*, *scopal specificity*, and *relational specificity*, as classified in von Heusinger (2002). The conclusion I arrive at is that *saai*'s specificity effect resembles neither of them, so the analyses for these types of specificity cannot be directly applied to account for the specificity effect of *saai*. I propose that *saai*'s specificity effect should be understood as a kind of specificity that targets distributive quantification. This type of specificity is referred to as independence in distributivity.

An epistemically specific indefinite refers to an individual that the speaker has in mind. Some studies take epistemic specificity to indicate that indefinites have a non-quantificational, referential use (Fodor and Sag 1982). von Heusinger (2002) uses the following example as an illustration.

(24) A student in Syntax 1 cheated on the exam.

a. His name is John.

b. We are all trying to figure out who it was.

In (24-a), the speaker can uniquely identify the individual the indefinite in (24) refers to, so the indefinite is said to be epistemically specific. However, if the speaker cannot uniquely identify the referent of the individual, as in (24-b), then the indefinite is said to be epistemically nonspecific.

When it comes to an indefinite following *saai*, it may be epistemically specific or not, as both types of follow-ups in (25-a) and (25-b) are felicitous. In other words, the specificity effect of *saai* is not epistemic specificity. If epistemic specificity is used to indicate that an indefinite has a referential use, then we can conclude that an indefinite following *saai* does not need to assume a referential use.

(25) Di-hoksaang gin-saai jat-go lousi.
CL.PL-student see-ASP one-CL teacher
'The students each saw a teacher.'

a. Zauhai Lei Lousi.
namely Lei Teacher
'Namely Teacher Lei.'

b. Dan ngo m-zi hai binggo lousi.
but I not-know be which teacher
'But I don't know which teacher.'

There are different definitions of scopal specificity in the literature, depending on whether non-island-bound scope (i.e., exceptional scope) is taken to be a defining feature. In this study, I follow von Heusinger (2002) and take scopal specificity to refer to the ability to take scope over another scope-bearing element (regardless of the presence of syntactic islands), such as negation, modals or conditionals.[7] *Saai* as a verbal suffix generally cannot attach to modal verbs, so the relevant

[7]Some studies take scopal specificity to follow from epistemic specificity (Fodor and Sag 1982) but others argue that it is a separate class (Farkas 1981, Kamp and Bende-Farkas 2019). Different theories of scopal specificity for indefinites have been developed over the years. When an indefinite does not occur in a syntactic island, the use of ordinary scope-taking suffices (Montague 1974, May 1977). When an indefinite occurs in an island, exceptional wide scope has been argued to come from (i) the use of (skolemized) choice functions (Reinhart 1997, Kratzer 1998, Matthewson 1999), (ii) anaphora to a quantificational structure (Brasoveanu and Farkas 2011, DeVries 2016), and (iii) the use of dynamic alternatives semantics (Charlow 2014). The analysis developed in this dissertation for *saai* is closest to the anaphora approach. I reserve a reincarnation of the present study in other frameworks for future research.

testing cases are negation and conditionals. von Heusinger (2002) cites the following example (from Karttunen 1976) to show the interaction of specificity and negation. To give rise to the interpretation in (26-a), the existential quantification is interpreted outside the scope of negation, so the indefinite is said to be scopally specific. By contrast, to give rise to the interpretation in (26-b), the existential quantification contributed by the indefinite is inside the scope of negation, so the indefinite is said to be scopally nonspecific.

(26) Bill didn't see a misprint.

   a. There is a misprint which Bill didn't see.

   b. Bill saw no misprints.

Similarly, indefinites in English may scopally interact with a conditional (Reinhart 1997). When it is interpreted as having wide scope relative to the conditional, as in (27-a), the interpretation is often said to be specific. When it is interpreted as having narrow scope relative to the condition, as in (27-b), the interpretation is referred to as non-specific.

(27) If a relative of mine dies, I will inherit a house.

   a. There is a particular relative of mine such that if s/he dies, I will inherit a house.

   b. If any relative of mine dies, I will inherit a house.

There are languages with morphology that gives rise to scopal specificity. For example, the determiner *ti* in St'át'imcets marks an indefinite that has to take wide scope relative to negation and conditional.

(28) cw7aoz kw-s áz'-en-as [ti sts'úqwaz'-a] kw-s Sophie
NEG DET-NOM buy-TR-3ERG DET fish-DET DET-NOM Sophie
'There is a fish which Sophie didn't buy.' (Matthewson 1999: (21))

(29) cuz' tsa7cw kw-s Mary lh-t'íq-as *ti* qelhmémen'-a
going.to happy DET-NOM Mary HYP-arrive-3CONJ DET old.person(DIMIN)-DET
'Mary will be happy if a particular elder comes.' (Matthewson 1999: (16))

However, a post-*saai* indefinite may scopally interact with negation, as already pointed out in (16) (repeated below):

(30) Di-hoksaang mou gin-saai jat-go lousi.
CL.PL-students not see-SAAI one-CL teacher
a. There is a teacher that the students did not all see.
b. There is no teacher that the students all saw.

Moreover, a post-*saai* indefinite may scopally interact with a conditional:

(31) Jyugwo di-jyu hai-saai jat-go gong japmin, ngo wui hou hoisam.
if CL.PL-fish in-SAAI one-CL tank inside I will very happy
a. There is a tank such that if all the fish are inside that tank, I'll be very happy.
b. If all the fish are in a single tank (regardless of which tank), I'll be very happy.

What this tells us is that the specificity effect of *saai* is not one that fully resembles scopal specificity. In particular, an indefinite following *saai* may freely interact with other scope bearing elements such as negation and conditionals. It looks, quite interestingly, that if any scopal effect is relevant, a post-*saai* indefinite is only required to be interpreted outside the scope of distributivity, not any other operator.

Finally, we test relative specificity. von Heusinger (2002), attributing the identification of this type of specificity to Enç (1991), uses the following example (due to Hintikka 1986) to illustrate it:

(32) According to Freud, every man unconsciously wants to marry a certain woman—his mother.

What is interesting is that there is no particular scope configuration of the indefinite that would give the sentence the intended interpretation. If the indefinite takes wide scope, we end up with a truth condition that is too strong: there is a particular woman that every man wants to marry. If the indefinite takes narrow scope, then the truth condition is too weak: as long as every man wants to marry some woman or other, the sentence is true. Rather, what the sentence requires is a specific relation linking the man and the woman—the woman is the man's mother.

Is the specificity effect of *saai* reducible to relational specificity? I think not. (33) cannot refer to a situation in which each fish is in a different tank even if the tanks happen to be the respective favorite tanks of the fish's. The only interpretation that is available is that the fish are all in the same tank and the tank is their favorite one.

(33) Di-jyu hai-saai jat-go gong japmin, jiuhai keoidei zeoi zungji-ge gong.
CL.PL-fish in-SIDE one-CL tank inside namely their most like-MOD tank.
'The fish are all in a tank, namely, their favorite tank.'

What is arising from this discussion is that the specificity effect of *saai* cannot be captured by epistemic specificity or relational specificity, and it only partially resembles scopal specificity. The kind of specificity effect we need is one that is intimately tied to distributivity. Let me suggest that this type of specificity be understood as **independence** relative to distributivity, formulated as a generalization below for easy reference:[8]

(34) The Independence Generalization of *saai*-distributivity

The evaluation of a post-*saai* expression is independent of the evaluation of distributive quantification.

All that *saai* requires is that an expression following it remain constant relative to distributive quantification. The expression may scopally interact with any other scope-bearing element as long as the independence generalization is satisfied. The rest of the chapter is devoted to two different accounts that derive the Independence Generalization. I ultimately argue in favor of the second account. However, exploring the first account offers some useful preparation for the second account.

The first account is outlined in section 4.3. According to this account, *saai* as a suffixal distributivity marker always combines with a verb and introduces distributivity that scopes only over the verbal predicate. A post-*saai* constituent, in this case, is naturally interpreted outside the scope of *saai*-distributivity.[9] I call this a **scope** account, for it derives the Independence Generalization by forcing distributivity contributed by *saai* to take narrow scope relative to post-*saai* nominals.

The second account is introduced in 4.4. In this account, *saai* is allowed to introduce distributivity that freely scopally interact with other scope expressions. However, a separate mechanism (formulated as a constraint) ensures that the post-*saai* constituent is interpreted *as if* it is outside

[8]I chose the term 'independence' rather than 'specificity' because the phenomena include both independence of individuals and independence of degrees (see Section 4.5). While independence of individuals can be think of as specificity, independence of degrees is much harder to think in terms of specificity.

[9]However, it may enter into scopal interactions with other sentential operators, and hence does not enjoy widest scope.

the scope of distributivity. I call this a **pseudo-scope** account. The pseudo-scope account is very similar to the scope account, but the there are empirical and conceptual differences that tell it apart from the former.

4.3 A scope account in terms of narrow-scope distributivity

The scope account relies on the key syntactic assumption that *saai* as a verbal suffix introduces distributivity that always takes narrow scope relative to other nominals. A concrete structural illustration is given in Figure 4.1. Since the scope account does not need to appeal to dynamic semantics or plural logic, I resume to a basic static semantics (with domain pluralities) throughout this subsection.

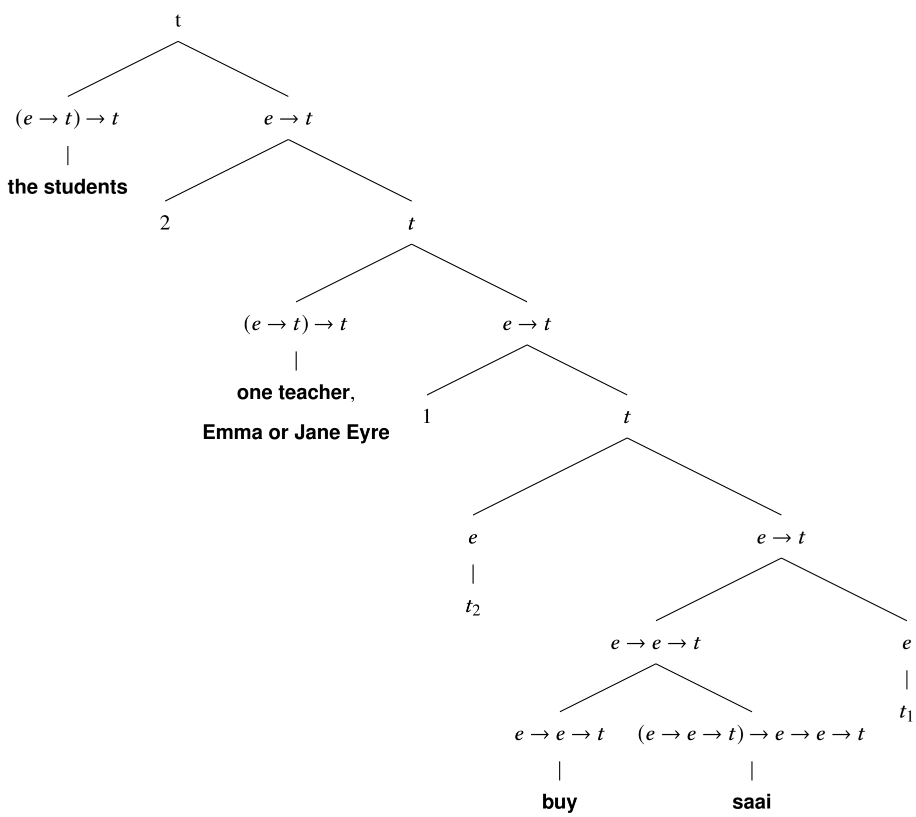


Figure 4.1: Narrow-scope distributivity

In this analysis, *saai* has the definition in (35). It takes a relation provided by a transitive verb, such as *see* in (111-a), and returns another relation. The newly returned relation is just like the original one except for the fact that the relation no longer holds between a subject and an object, but between the atomic parts of the subject and an object, as shown in (37).

(35) **saai** := $\lambda R \lambda y \lambda x.\forall z[z \leq_A x](R\ \ y\ z)$, where $\leq_A$ is the 'atomic part-of' relation

(36) **gin** 'see' := $\lambda y \lambda x.\textsf{see}\ y\ x$

(37) **gin-saai** := $\lambda y \lambda x.\forall z[z \leq_A x](\textsf{see}\ \ y\ z)$

Saai is flexible with the arity of its relational argument, as it can also combine with an intransitive verb, as shown in (38).

(38) Keoidei zau-saai la.
they *leave*-SAAI SFP
'They each left.'

For this reason, *saai*'s argument structure should be generalized. Instead of taking a relation and returning a relation

a two-place relation ($e \to e \to t$) or a relation with a higher arity). A type-flexible definition of *saai* is offered below:

(39) **saai** := $\lambda \alpha \lambda \vec{y} \lambda x.\forall z[z \leq_A x](\alpha\ \ y\ z)$, where $\leq_A$ is the 'atomic part-of' relation, α a n-nary predicate, and $\vec{y}$ is a sequence of $n-1$ variables.

4.3.1 Cardinal indefinites

A cardinal indefinite denotes a generalized quantifier, as shown in (40) (e.g., Montague 1974, Barwise 1981). Following Montague (1974), a plural definite subject can also be modeled as a generalized quantifier, as shown in (41).

(40) **jat-go lousi** 'one teacher' := $\lambda P.\exists y[\textsf{book}\ y\ \wedge\ \mu_{\text{CARD}}\ y = 1](P\ y)$

(41) **Di-hoksaang** 'the students' := $\lambda P.P(\bigoplus \textsf{stdts})$

Folding in the lexical ingredients in (37), (40) and (41), a sentence with a cardinal indefinite following *saai* is then interpreted as follows:

(42) a. **the students** $(\lambda x.$ **one teacher** $(\lambda y.$ **see-saai** $y\ x))$

b. $\exists y[\text{book } y \wedge \mu_{\text{CARD}} = 1](\forall z[z \leq_A \bigoplus \text{stdts}](\text{read } y\ z))$

The cardinal indefinite naturally takes wide scope over the distributive quantification introduced by *saai*. This is because the distributive quantification is introduced, in the first place, as only having scope over an individual, i.e., the 'trace' of a quantificational object. So, it is not surprising that the cardinal indefinite ends up taking wider scope over distributivity. This also essentially ensures that the cardinal indefinite is interpreted as independent of, i.e, not co-vary with, distributivity.

For ease of comparison, let me illustrate the range of possible interpretations for a sentence with a cardinal indefinite and a distributivity marker like *cyubou* or *dou*. Assume that these distributivity markers contribute a standard VP-level distributivity operator, which can scopally interact with cardinal indefinites. The scope interaction then gives rise to two interpretations: the "distributivity > indefinite" interpretation in (43) and the "indefinite > distributivity" interpretation in (44).

(43) a. **the students** (**cyunbou/dou** $(\lambda x.$ **one teacher** $(\lambda y.$ **see** $y\ x)))$

b. $\forall z[z \leq_A \bigoplus \text{stdts}](\exists y[\text{teacher } y \wedge \mu_{\text{CARD}} = 1](\text{see } y\ z))$

(44) a. **one teacher** $(\lambda y.$ **the students** (**cyunbou/dou** $(\lambda x.$ **see** $y\ x)))$

b. $\exists y[\text{teacher } y \wedge \mu_{\text{CARD}} = 1](\forall z[z \leq_A \bigoplus \text{stdts}](\text{see } y\ z))$

4.3.2 Disjunction

Disjunction has long been noted to participate in scopal interactions (e.g., Larson 1985). Since proper names can be lifted to generalized quantifiers (Partee 1986; see also Montague 1974), a disjunction involving two proper names can be treated as a disjoined generalized quantifier following Rooth and Partee (1982), as shown in (45). This generalized quantifier occupies the same position as a cardinal indefinite. For this reason, a sentence with a disjunction following *saai* also naturally has the disjunction out-scoping the distributivity, as shown in the LF in (46-a) and the semantic translation in (46-b).

(45) **Emma or Jane Eyre** $:= \lambda P.P$ **e** $\vee$ P **je**

(46) a. **the students** $(\lambda x.$ **Emma or Jane Eyre** $(\lambda y.$ **read-saai** $y\ x))$

b. $\forall z[z \leq_A \bigoplus \text{stdts}](\text{read e } z) \vee \forall z[z \leq_A \bigoplus \text{stdts}](\text{read je } z)$

If distributivity is introduced not by *saai* but by an adverbial distributivity marker capable of scopally interacting with disjunction, such as *cyunbou* or *dou*, then the corresponding sentence is ambiguous, as we have seen in the case of cardinal indefinites.

(47) a. **the students** (**cyunbou/dou** (λx. **Emma or Jane Eyre** (λy. **read** $y\ x$)))

b. $\forall z[z \leq_A \bigoplus \text{stdts}](\text{read je } z \vee \text{read e } z)$

(48) a. **Emma or Jane Eyre** (λy. **the students** (**cyunbou/dou** (λx. **read** $y\ x$)))

b. $\forall z[z \leq_A \bigoplus \text{stdts}](\text{read e } z) \vee \forall z[z \leq_A \bigoplus \text{stdts}](\text{read je } z)$

4.3.3 Bare noun phrases

Treating bare NPs requires some caution. If we assume that bare NPs are existential quantifiers like cardinal indefinites, then the prediction is that they pattern like cardinal indefinites in their interactions with *saai*. This is not in accordance with the empirical generalization. As reported in Section 3.2, bare noun phrases are allowed to have witnesses that co-vary with distributivity. To model the interactions between bare noun phrases and *saai*, I suggest we exploit a longstanding tradition in semantics to treat bare noun phrases as proper names of kinds (Carlson 1977a,b, Chierchia 1998, Dayal 2004, 2011a). On this view, they are not scope-bearing elements and may directly serve as an argument for a predicate that has composed with *saai*.[10] As an example, the bare noun phrase *syu* 'book(s)' is translated and abbreviated as follows[11]:

(49) **syu** 'book(s)' $:= \lambda s.\iota x.\text{bks}_s\ x$

$= \text{bk-kind}$

Plugging in this bare noun phrase into the structure in Figure 4.1 yields the LF (50-a) and its semantic translation in (50-b).

(50) a. **the students** (λx.**see-saai book(s)** x)

[10] Even if they are lifted to generalized quantifiers, they behave like proper names and do not enter into scope interactions with other operators.

[11] Bare noun phrases are number-neutral in Cantonese.

b. $\forall z[z \leq_A \bigoplus \text{stdts}](\text{read bk-kind } z))$

It is well known in the literature of bare noun phrases that a sortal repair strategy is needed to compose an object-level predicate (like *read*) and a kind term contributed by a bare noun phrase (Carlson 1977a,b, Chierchia 1998, Dayal 2004, 2011a). In this study, I adopt Derived Kind Predication (DKP), as proposed in Chierchia (1998) to repair the sortal mismatch. This sortal repair strategy is defined as follows ($^{\cup}$ shifts a kind to a property)[12]:

(51) **DKP**

If R is an n-place relation over individuals and k a kind term, then:

$R(k) := \lambda x_1, ..., \lambda x_{n-1}.\exists y[^{\cup}k\ y](R\ y\ x_1, ..., x_{n-1})$

Note that the existential quantification introduced by the sortal repair strategy always takes the narrowest scope (Carlson 1977a,b, Chierchia 1998). Therefore, applying DKP to (50-b) yields a narrow scope existential interpretation, as shown in (52), which is compatible with witness variation.

(52) $\forall z[z \leq \bigoplus \text{stdts}](\exists y[^{\cup}\text{bk-kind } y](\text{read } y\ z))$

In short, according to this analysis, *saai* contributes narrow-scope distributive quantification. As a result, the distributive quantification fails to interact with other scopal expressions, such as cardinal indefinites and disjunction, giving rise to a wide-scope interpretation of these scopal expressions. Bare noun phrases are exceptions because they induce existential quantification as a sortal repair strategy, which always takes the narrowest scope.

4.3.4 Multiple post-*saai* constituents

A very nice prediction of the narrow scope distributivity account is that any cardinal indefinite introduced following *saai* has to not co-vary with distributivity. This is because *saai* is stipulated to introduce distributivity scoping only over the verbal relation, hence any number of cardinal indefinites (or disjunction) should be interpreted outside the scope of distributivity. This prediction is borne out by the following examples:

[12]This DKP is a relational version suggested in Chierchia (1998:fn.16)

(53) Keoidei song-saai jat-bun syu bei jat-go hoksaang
they give-SAAI one-CL book to one-CL student
'They all gave a particular book to a particular student.'

(54) Keoidei giu-saai jat-go jan heoi maai jat-bun syu
they ask-SAAI one-CL person buy one-CL book
'They all asked a particular person to buy a particular book.'

The narrow-scope distributivity analysis is straightforward and accounts for the data set introduced in section 3.2. However, it runs into a few empirical issues, which are discussed in the following subsections.

4.3.5 Empirical problem 1: cardinal indefinites with bound pronouns

When a cardinal indefinite contains a pronoun in the common noun restriction bound by the distributivity key, as shown in (96) (the entire cardinal indefinite is enclosed in "[...]"), the cardinal indefinite can co-vary with the distributivity key:

(55) Di-hoksaang^x maai-saai [*zigei*$_x$ jungji-ge jat-bun syu].
CL-PL-students buy-SAAI self like-MOD one-CL book
'The students bought a book they like.'

The behavior of this type of cardinal indefinites cannot be accounted for by simply letting the cardinal indefinites be interpreted outside the scope of distributivity. To see this, I first translate the cardinal indefinite as an existential quantifier with the pronoun interpreted as a free variable.

(56) **zigei**$_x$ **jungji-ge jat-bun syu** 'one book self$_x$ like'

$$:= \lambda P.\exists y[\textsf{book}\ y\ \wedge\ \mu_{\text{CARD}}\ y = 1\ \wedge\ \textsf{like}\ y\ x](P\ y)$$

After plugging the indefinite into the structure in Figure 4.1, we obtain the LF in (57-a) and its semantic translation in (57-b). However, this interpretation does not allow different students buying different books. What it allows is every student reading a single book that they collectively like. Even if we assume that 'collective liking' gets resolved in the same way as distributive liking, (57-a) and (57-b) are still inadequate because they do not allow the books to co-vary with the students.

(57) a. **the students** (λx. **one book self**$_x$ **like** (λy. **buy-saai** y x))

b. $\exists y[\text{book } y \wedge \mu_{\text{CARD}}\ y = 1 \wedge \text{like } y \bigoplus \text{stdts}](\forall z[z \leq_A \bigoplus \text{stdts}](\text{buy } y\ z))$

The behavior of indefinites with bound pronouns can be accounted for *if* we assume that an indefinite with a bound pronoun falls inside the scope of the distributive quantification introduced by *saai*. This is because the pronoun will be bound by the universal quantifier that quantifies over the atomic parts of the plurality denoted by the distributivity key. The corresponding semantic interpretation is given below:

(58) $\forall z[z \leq \bigoplus \text{stdts}](\exists y[\text{book } y \wedge \mu_{\text{CARD}}\ y = 1 \wedge \text{like } y\ z](\text{read } y\ z))$

However, there is no way to sneak a cardinal indefinite back into the scope of the distributive quantification introduced by *saai*, given that the distributive quantification is formulated to only scope over individuals, i.e., which are traces of quantifiers like cardinal indefinites. If we are to assume that *saai* has an alternative lexical entry allowing the distributive quantification it introduces to scopally interact with cardinal indefinites with a bound pronoun, we then need to justify what bans this lexical entry in the cases of simple cardinal indefinites and disjunction.

In short, a cardinal indefinite with a bound pronoun imposes conflicting requirements on the relative scope of the cardinal indefinite and the distributive quantification introduced by *saai*: to maintain the integrity of the narrow-scope distributivity analysis, the indefinite should be interpreted *outside* the scope of distributivity; however, to model the co-variation induced by the bound pronoun, the indefinite has to be interpreted *inside* the scope of distributivity. Without further assumptions, it is not clear how an indefinite can be inside and outside the scope of distributivity at the same time.

4.3.6 Empirical Problem 2: scope interference with other distributivity markers

There are two other distributivity markers in Cantonese: *cyunbou* (59-a) and *dou* (60-a). *Saai* can co-occur with both without inducing ungrammaticality, as evidenced by (59-b) and (60-b).

(59) a. Di-hoksaang cyunbou gin-zo Can lousi.
CL.PL-students all see-ASP Can teacher
'The students all saw Teacher Can.'

b. Di-hoksaang cyunbou gin-saai Can lousi.
CL.PL-students all see-SAAI Can teacher
'The students all saw Teacher Can.'

(60) a. Cyun-ban hoksaang dou gin-zo Can lousi.
whole-class student DOU see-ASP Can teacher
'Each student in the class met Teacher Can.'

b. Cyun-ban hoksaang dou gin-*saai* Can lousi.
whole-class student DOU see-SAAI Can teacher
'Each student in the class met Teacher Can.'

There are two reasons why co-occurring distributivity markers are of interest to this study. First, if *saai*, *cyunbou* and *dou* all contribute genuine distributive quantification targeting the same plural subject, it is unclear why they do not give rise to vacuous distributive quantification, which is banned in languages like English.[13]

(61) *Every student each saw Miss Carla.

Admittedly, co-occurring distributivity markers is a poorly understood phenomenon. What it challenges is the practice of translating every instance of markers of distributivity as an independent distributivity operator, rather than the proposal that *saai* contributes narrow-scope distributivity.[14]

For this reason, it is useful to consider another interesting pattern resulting from co-occurring distributivity: scope interference. Concretely, when *cyunbou* and *dou* occur without *saai*, they are capable of scopally interacting with an indefinite, as pointed out in Section 3.2. For example, both (62) and (63) allow a narrow-scope interpretation of the cardinal indefinite in the object position.

(62) Di-hoksaang cyunbou gin-zo jat-go lousi.
CL.PL-students all see-ASP one-CL teacher

[13] Many languages allow more than one distributivity marker to occur in a sentence without inducing double distributivity. For example, Kaqchikel (Henderson 2014), Hungarian (Kuhn 2017), and American Sign Language (Kuhn 2017) allow the co-occurrence of a distributive quantifier and a distributive numeral. Relatedly, Szabolcsi (2010) reports that to some (but not all) speakers of English, distributive quantifiers may co-occur with binominal *each*, as in (i) (cited from Kuhn 2017: (15)).

(i) %Every job candidate was in the room for fifteen minutes each.

[14] That being said, the analysis developed in Section 4.4 does address the co-occurrence puzzle of distributivity markers to some extent. The spirit of the analysis is that *saai* does not in fact introduce distributive quantification. Rather, it imposes an 'independence constraint' on the functional dependency resulting from distributive quantification.

‘The students each saw a teacher (possibly different ones).’

(63) Cyun-ban hoksaang dou gin-zo jat-go lousi.
whole-class student DOU see-ASP one-CL lousi
‘Each student in the class met a teacher (possibly different ones).’

However, when *saai* surfaces in these sentences, *cyunbou* and *dou* fail to scopally interact with the cardinal indefinites, as shown in (64) and (65). It is as though *saai*’s presence interferes with the scope interactions between *cyunbou*/*dou* and other scopal expressions.

(64) Di-hoksaang cyunbou gin-saai jat-go lousi.
CL.PL-students all see-SAAI one-CL teacher
‘The students each saw a teacher (the same teacher).’

(65) Cyun-ban hoksaang dou gin-*saai* jat-go lousi.
whole-class student DOU see-SAAI one-CL lousi
‘Each student in the class met a teacher (the same teacher).’

Treating *saai* as merely contributing narrow-scope distributivity does not account for its ability to induce scope interference.

4.3.7 Interim summary

Given the empirical challenges faced by the narrow-scope distributivity account, I do not think it holds the ultimate key to analyzing *saai* distributivity. However, there is no denying that interpreting indefinites and disjunction outside the scope of distributivity does offer a relatively natural and simple analysis for their lack of co-variation with distributive quantification. In the formulation of an alternative analysis to address the empirical issues, it is worth preserving the simplicity of the narrow-scope distributivity account.

In the next section, I offer a pseudo-scope account that mimics the narrow-scope distributivity account very closely in terms of the predictions for cardinal indefinites, disjunction, and bare noun phrases in *saai*-distributivity. The account crucially relies on the use of an **independence constraint**, imposed on the functional dependency arising from distributive quantification. To model the fact that distributive quantification contributes a functional dependency that can be subject to further constraints, I use the framework developed in Chapter 2.

4.4 A pseudo-scope account in the framework of DPlLM

4.4.1 Proposal: an independence constraint

The pseudo-scope account has the following main ingredients:

- *Saai* is not treated as distributivity operator. Rather, I argue that it imposes a constraint on the functional dependency arising from distributive quantification.
- Distributivity, as contributed by *cyunbou*, *dou*, or a null distributivity operator, is allowed to freely scopally interact with any post-*saai* scopal expressions. In other words, the assumption that *saai* contributes narrow-scope distributivity is removed.
- The constraint contributed by *saai* requires that **values introduced inside the scope of distributivity by a post-*saai* expression remains independent of distributivity**. More precisely,
  - if a post-*saai* expression is interpreted outside the scope of distributive quantification, nothing happens to it; however,
  - if a post-*saai* expression is interpreted inside the scope of distributive quantification, then it is required to have a *constant witness* relative to, i.e., not co-vary with, distributive quantification.

The last point is particularly important. It amounts to giving an expression inside the scope of distributive quantification **pseudo wide-scope**. In fact, quite a number of indefinites have received a pseudo-scope account using choice functions, such as indefinites marked by *(t)i-* in Sta'át'imcets (Matthewson 1999) and indefinites marked by the suffix *-khí* in Tiwa (Dawson 2018).

While analyzing indefinites marked by *saai* as choice function indefinites seems like a plausible option, it does not account for why the wide-scope behavior of *saai* is inherently tied to distributivity.[15] In other words, *saai* does not mark the wide scope status of an indefinite when there is no distributivity, as shown in (66), unlike the wide-scope markers in Sta'át'imcets (67) or Tiwa (68).

[15] For additional challenges faced by the choice-function approach to indefinites, see Geurts (2000).

(66) #Jyugwo lei gin-saai jat-go hoksaang, ngo wui hou hoisam.
if you see-SAAI one-CL student I will very happy
Intended 'If you see a particular student, I will be very happy.'

(67) cuz' tsa7cw kw-s Mary lh-t'íq-as *ti* qelhmémen'-a
going.to happy DET-NOM Mary HYP-arrive-3CONJ DET old.person(DIMIN)-DET
'Mary will be happy if a particular elder comes.' (Sta'át'imcets, Matthewson 1999: (16))

(68) Maria inda-*khí* kashóng pre-ya-m.
Maria what-KHI dress buy-NEG-PST
'Maria didn't buy some dress.' (Tiwa, Dawson 2018: (20))

In this paper, I pursue a pseudo-scope account couched in the framework of Dynamic Plural Logic with Measurement. The primary reason for using this logic is because it allows distributive quantification to contribute functional dependencies that can be passed down from context to context. In other words, distributivity is fully dynamic in this logic. For this reason, we can talk not only about *introducing* distributivity into context, but also *retrieving* it from context. The latter is an important component in the semantics of *saai*, which is analyzed as imposing a constraint on the distributivity dependency it accesses anaphorically.

Using DPlLM to model the independence constraint of *saai* has an extra benefit: it allows us to directly compare *saai* and distributive numeral markers, which have been analyzed using PCDRT, a cousin logic of DPlLM. The comparison is offered in the last part of this section.

At the core of *saai*'s contribution is an **independence constraint**, formulated as in (69).

(69) $G[\![\mathsf{ind}_{x,\vec{y}}]\!]H := G = H$ and for all $a, b \in G(x) : G|_{x=a}(\vec{y}) = G|_{x=b}(\vec{y})$
'y's value is constant relative to x's value.'

Saai anaphorically accesses the functional dependency introduced by distributive quantification with help of the first index, i.e., the variable x that stores the values contributed by the distributivity key. Then, it accesses all the new variables $y_1, y_2, ...y_n$ introduced *in the scope of distributive quantification*, and requires that each variable stores values that are constant relative to the values in the distributivity key. In other words, although distributive quantification allows a variable introduced in its scope to exhibit dependence with the variable storing the values associated with the distributivity key, *saai* effectively forbids the dependence. What this amounts to is that the variable is introduced *as if* it is outside the scope of distributivity. A very similar account is offered in Brasoveanu and Farkas (2011) for indefinites that seem to take exceptional scope. This is how the independence constraint mimics the scope account.

4.4.2 Proper names, definite descriptions, Cardinal indefinites, and disjunction

In this section, I discuss four types of expressions that are interpreted as independent of the distributivity key. The first two types, proper names and definite expressions, naturally do not induce variation in the scope of distributive quantification. So, the independence constraint, when applied to them, is trivially satisfied. The remaining two types of expressions, namely, cardinal indefinites and disjunction, may induce variation in the scope of distributive quantification. However, the independence constraint forces them to lack co-variation.

A notational note before proceeding. To differentiate between the static semantics used in the narrow scope distributivity analysis and the constraint-based analysis, slightly different symbols are used to translate lexical entries and phrases in the two types of semantics.

Proper names and definite expressions Proper names and definite expressions may not co-vary with distributivity, so the independence constraint has no effect on them. In particular, both types of expressions introduce into an info-state a variable storying a fixed set of values that do not change in the course of evaluating distributive quantification. This can be shown with the definition of the proper name *Mingzai* in (70-a) and the definition of the definite expression *di-lousi* 'the teachers' in (71-a).

(70) a. $\textbf{Mingzai}^{y} := \lambda P.\exists y \ \wedge y = \mathsf{m} \ \wedge P\ y$

b. $\mathsf{max}^{x}(\mathsf{stdts}\ x) \wedge \delta_x(\exists y \wedge y = \mathsf{m} \wedge \mathsf{see}\ y\ x) \,\overline{\wedge}\, \mathsf{ind}_{x,y}$

(71) a. $\textbf{the teachers}^{y} := \lambda P.\mathsf{max}^{y}(\mathsf{teachers}\ y)\ \wedge P\ y$

b. $\mathsf{max}^{x}(\mathsf{stdts}\ x) \wedge \delta_x(\mathsf{max}^{y}(\mathsf{teachers}\ y)\ \wedge \mathsf{see}\ y\ x) \,\overline{\wedge}\, \mathsf{ind}_{x,y}$

In (70-a), variable introduction introduces the variable y and $y = \mathsf{m}$ ensures that y is associated with only one value, namely, m. Evaluating $\exists y$ in the scope of the distributivity operator δ_x in (70-b) may give rise to covariation between the students stored in x and the random values stored in y, but the next conjunct $y = \mathsf{m}$ makes sure to remove all the info-states in which y has any value other than m. This essentially ensures that for all values in x, the corresponding y-value can only be m, i.e., the individual Mingzai. Similarly, in (71-a), maximization over y makes sure that y stores all the teacher values in the model. So, even if the maximization over y falls inside the scope of the distributivity operator δ_x in (71-b), there is no co-variation between the students and the teachers they saw. Therefore, when a post-*saai* expression is a proper name or a definite expression, the independence constraint is guaranteed to be satisfied as long as the distributivity contribution is true.

<u>Cardinal indefinites</u> Cardinal indefinites are translated as dynamic generalized quantifiers, as shown in (72), with $\exists y$ understood as variable introduction.

(72) **jat-go lousi** 'one teacher' := $\lambda P.\exists y \wedge \mathsf{book}\ y \wedge \mu_{\text{CARD}}\ y = 1 \wedge P\ y$

A definite plural like *di-hoksaang* 'the students' is also treated as a dynamic generalized quantifier, as shown in (73). It introduces into an info-state a d-ref associated with the maximal plural individual that satisfies the common noun restriction. The maximal plural individual is obtained in the manner stated in (74) (see Chapter **?** for how maximization works).

(73) **Di-hoksaang** 'the students' := $\lambda P.\mathsf{max}^y(\mathsf{stdts}\ y) \wedge P\ y$

(74) $G[\![\mathsf{max}^x(P\ x)]\!]H = \mathbb{T}$ iff $G[\![\exists x \wedge P\ x]\!]H = \mathbb{T}$ and there is no H', such that $\bigoplus H'(x) > \bigoplus H(x)$ and $G[\![\exists x \wedge P\ x]\!]H' = \mathbb{T}$

To see how the independence constraint contributed by *saai* constrains the scope interaction between a distributivity operator and a cardinal indefinite, let us consider a concrete sentence with three elements: (i) distributive quantification induced by *cyunbou* or *dou*, (ii) a cardinal indefinite, and (iii) *saai*.

As discussed in Section 4.3, the distributivity operator and the cardinal indefinite may enter into scope interactions. For our purpose, let us first zoom into the LF in (75-a), in which the distributivity operator takes wide scope over the cardinal indefinite. The resulting interpretation is given in (75-b).

(75) a. **the students**^{x} (**cyunbou/dou**$_x$ (λu. **one book**^{y} ($\lambda u'$. **read** u' u)))

b. $\mathsf{max}^x(\mathsf{stdts}\ x) \wedge \delta_x(\exists y\ \wedge\ \mathsf{book}\ y\ \wedge \mu_{\text{CARD}}\ y = 1\ \wedge\ \mathsf{buy}\ y\ x)$

Evaluation of such a formula in DPlLM against an input info-state gives rise to a set of info-states. Suppose H and H' in Figure 4.2 are two info-states in the output set. In both info-states, x stores the student values contributed by the distributivity key, and y stores a set of book values contributed by the cardinal indefinite. In addition, each x-value bought the corresponding y-value as instructed by the assignment functions. The two info-states differ in the values associated with d-ref y introduced by the cardinal indefinite. In H, y stores a singleton set of values, while in H', y stores a set of three values.

H	x	y
h_1	s1	bk1
h_2	s2	bk1
h_3	s3	bk1

Info-state H

$H|_{x=\mathsf{s1}}(y) = \{\mathsf{bk1}\}$

$H|_{x=\mathsf{s2}}(y) = \{\mathsf{bk1}\}$

$H|_{x=\mathsf{s3}}(y) = \{\mathsf{bk1}\}$

H'	x	y
h'_1	s1	bk1
h'_2	s2	bk2
h'_3	s3	bk3

Info-state H'

$H'|_{x=\mathsf{s1}}(y) = \{\mathsf{bk1}\}$

$H'|_{x=\mathsf{s2}}(y) = \{\mathsf{bk2}\}$

$H'|_{x=\mathsf{s3}}(y) = \{\mathsf{bk3}\}$

Figure 4.2: H satisfies the independence constraint of *saai* while H' does not

Now, we are ready to add the contribution of *saai*. *Saai*, together with the d-refs it anaphorically accesses, contributes an independence constraint, i.e., the last conjunct in (76). Recall that evaluation of the first two conjuncts in (76) returns the set of info-states in Figure 4.2. The independence constraint is imposed on this output set. For each info-state in Figure 4.2, we check to see if it satisfies the independence constraint. If it does, the info-state is kept; otherwise, it is discarded. If the after evaluating the constraint the final output has at least one info-state, the sentence is true. If the final output is an empty set, the sentence is not false, but undefined (see the definition of "$\overline{\wedge}$" in (99)).

(76) $\underbrace{\textsf{max}^x(\textsf{stdts}\ x) \wedge \delta_x(\exists y \wedge \textsf{book}\ y\ \wedge \mu_{\text{CARD}}\ y = 1 \wedge \textsf{buy}\ y\ x)}_{\text{Distributive quantification}}\ \overline{\wedge}\ \underbrace{\textsf{ind}_{x,y}}_{\text{Constraint}}$

The independence constraint checks the values stored in two d-refs. The firs d-ref is the variable storing the values associated with the distributivity key. The second d-ref is the variable introduced inside the distributive scope of the firs variable that stores the values introduced by a post-*saai* expression. The constraint requires that the latter be independent of the former. When the independence constraint is imposed on H, it is satisfied. However, when the same constraint is imposed on H', it is not satisfied. This is because not all x-values are associated with the same y-value in H'. As long as there is an info-state like H in the output that can satisfy the independence constraint, the sentence with *saai* followed by a cardinal indefinite is judged to be true.

We have seen what happens when a cardinal indefinite is interpreted *inside* the scope of distributivity. Now, what happens when a cardinal indefinite is interpreted *outside* the scope of distributive quantification, as indicated in the LF in (77-a) and the formula in (77-b)?

(77) a. **the students**^{x} (λu. **one book**^{y} ($\lambda u'$. **cyunbou/dou**$_x$ (u bought u')).

b. $\textsf{max}^x(\textsf{stdts}\ x)\ \wedge\ \exists y \wedge \textsf{book}\ y\ \wedge \mu_{\text{CARD}}\ y = 1 \wedge\ \delta_x(\textsf{buy}\ y\ x)$

Nothing in the pseudo-scope account so far says anything about this scope configuration. However, I assume that *saai* does not track the independence of a d-ref if it is introduced outside the scope of distributivity. This assumption has a few merits. First, it is empirically more adequate. As will be seen in the discussion of indefinites with a bound pronoun, expressions capable of establishing dependency without being inside a distributivity operator are exempt from the independence constraint. Second, this assumption is closely related to how indefinites receive its interpretation in Brasoveanu and Farkas (2011). In their study, an indefinite chooses its (in)dependence, by tracking and relating to variables introduced by structurally more dominant quantifiers. In this study, a distributivity marker chooses the (in)dependence of an indefinite, by tracking and relating to variables dominated by them, i.e., variables introduced in their scope.

Disjunction The same account can be straightforwardly extended to disjunction, which also scopally interacts with distributivity. First, let us define a disjunctive DP in DPlLM:

(78) **Emma or^y Jane Eyre** $:= \lambda P.\exists y \wedge (y = \mathsf{e} \vee y = \mathsf{je}) \wedge P\, y$

A disjunctive DP is a dynamic generalized quantifier, just like an ordinary, non-disjunctive DP. The only difference is that it introduces a d-ref that stores values corresponding to either of the disjoined DPs. For simplicity, I only illustrate the interpretation when disjunction takes narrow scope relative to distributivity, as indicated in the LF in (79-a) and the formula in (79-b).

(79) a. **the students^x** (**cyunbou/dou** (λu. **Emma or^y Jane Eyre** ($\lambda u'$. **read** $u'\ u$)))

b. $\mathsf{max}^x(\mathsf{stdts}\ x) \wedge \delta_x(\exists y \wedge (y = \mathsf{e} \vee y = \mathsf{je}) \wedge \mathsf{read}\ y\ x)\ \overline{\wedge}\ \mathsf{ind}_{x,y}$

The independence constraint forces the disjunction to be associated with a set of values fixed relative to the distributivity key.

4.4.3 Bare noun phrases

Bare noun phrases in a post-*saai* position have interesting properties. First, unlike cardinal indefinites, bare noun phrases are not marked. More importantly, bare noun phrases are allowed to co-vary with distributivity. In other words, they seem to be immune to the independence constraint. The following examples help illustrate these two properties:

(80) Di-hoksaang maai-saai syu.
CL.PL-student buy-SAAI book
'The students each bought one or more books (possibly different ones).'

(81) Di-jyu hai-saai gong japmin.
CL.PL-fish in-SAAI tank inside
'The fish each are in a tank (possibly different ones).'

Since Carlson (1977a,b), it is widely recognized that ordinary indefinites and bare noun phrases are semantically quite different. So, it is not entirely surprising that bare noun phrases do not pattern like cardinal indefinites in Cantonese with respect to the independence constraint. That said, it is still desirable to have a concrete way to model the differences between bare noun phrases and cardinal indefinites that are responsible for the distinct interactions with the independence constraint. I explore a possibility below.[16]

I propose that a bare noun phrase is like a proper name for the purpose of the independence constraint. Carlson (1977a,b) has explicitly argued to treat bare plurals in English as proper names of kinds. This analysis has been extended to bare noun phrases in Cantonese by Cheng and Sybesma (1999) and Jiang (2012). Following their analysis, a bare noun phrase like *syu* 'book(s)' in Cantonese can be translated as follows into DPlLM:

(82) $\textbf{book(s)} := \lambda P.\exists y^k \wedge y^k = \text{book-kind} \wedge P\ y^k$

A bare noun phrase is very similar to a proper name, with the exception that the d-ref being introduced is a kind-level d-ref. To distinguish between an individual-level d-ref as well as a kind-level d-ref, I notate the latter as x^k. This notation is borrowed from the literature of kinds terms (Carlson 1977a,b, Yang 2001, Dayal 2011a). Carlson (1977b) and Dayal (1999) have demonstrated that pronominal anaphora to kinds are acceptable in English and Hindi. (83) shows that pronominal anaphora to a kind is also possible in Cantonese.

(83) Ngo cammaan gin-dou songsyu$^{x^k}$. Ngo zidou nei zungji keoidei$_{x^k}$, soyi ngo jing-zo
I last.night see-ASP squirrels I know you like them, so I take-ASP
zeong seong bei nei taai.
CL picture to you see
'I saw squirrels last night. I know you like them, so I took a picture for you.'

The intended interpretation for (83) is for the plural pronoun to refer to squirrels in general rather

[16] Another possibility I have not explored is to treat bare noun phrases in Cantonese as semantically incorporated and do not introduce d-refs (see also Dayal 1999, 2011b, Farkas and de Swart 2003, Krifka and Modarresi 2016).

than to the particular squirrels the speaker saw. The fact that this type of anaphora is possible indicates that the antecedent bare noun phrase *songsyu* 'squirrel(s)' introduces a kind-level d-ref that can be anaphorically accessed later by a plural pronoun.

A immediately merit of the kind-based analysis is that it allows the independence constraint to be satisfied with use of the kind-level d-ref. This is shown in (84). After all, the kind-level d-ref is just like a d-ref storing a proper name, which does not vary with distributivity.

(84) a. **the students**^{x} (**cyunbou/dou** (λu.. **read books**$^{y^k}$ u)))

b. $\mathsf{max}^x(\mathsf{stdts}\ x) \wedge \delta_x(\exists y^k \wedge y^k = \mathsf{bk\text{-}kind} \wedge \mathsf{buy}\ y^k\ x) \overline{\wedge}\, \mathsf{ind}_{x,y^k}$

There are two challenges for only recognizing the kind-level contribution of a bare noun phrase. The first one is that an additional mechanism is needed to evaluate a lexical relation involving a kind-level d-ref. In other words, a mechanism is needed for properly interpreting $\mathsf{buy}\ y^k\ x$ in (84). The literature has offered a few sortal repair strategies, including the stage predication proposed in Carlson (1977b) and Derived Kind Predication proposed in Chierchia (1998) (cf. Dayal (2013)). Despite their differences, they both share the effect of turning a kind into a (possibly plural) concrete individual. Since we have seen Chierchia (1998)'s Derived Kind Predicate in the discussion of bare noun phrases in Section 4.3, I use it as a basis for formulating a special evaluation rule involving an individual-level relation and a kind-level d-ref:

(85) **DKP** (in DPlLM)

If R is an 2-place relation over individuals and y^k a kind-level variable, then:

$G[\![R\ y^k\ x]\!]H = \mathbb{T}$ iff $G = H$ and there is a possibly plural z such that $z \in {}^{\cup}\bigoplus[\![y^k]\!]^G$ and $\langle \bigoplus[\![x]\!]^G,\ z\rangle \in \mathcal{I}(R)$

Formulating DKP as an evaluation rule linking a dynamic proposition to its truth condition essentially makes it a static procedure. The existential quantification over individuals in the instantiation set of the kind is an ordinary static existential quantifier and cannot introduce discourse referents into an info-state. Without a d-ref storing the individuals (in addition to the kind) it is harder to model anaphoric reference to individuals. As shown in (86), anaphoric reference to individuals appears to be possible with a bare noun phrase antecedent in Cantonese. I do not have a satisfactory account for how to model pronominal anaphora involving bare noun phrases. Some proposals targeting this phenomenon have been developed by Dayal (2011b) and Krifka and Modarresi (2016).

(86) Mingzai hai gongjuan gindou siupangjau. Keoidei wan-dak hou hoisam.
Mingzai in park see child they play-RES very happy
'Mingzai saw some children in the park. They were playing happily.'

Before ending this section, I would like to note a difference between indefinites and bare noun phrases that is indicative of their different discourse status (p.c. Simon Charlow and Veneeta Dayal). The difference lies in the so-called uniqueness implication. A hallmark property of anaphora involving indefinites is that it may (but not necessarily) lack a uniqueness implication (Heim 1982, 1990; cf. Evans 1977). An example showing this is given below:

(87) There once was a doctor in London. He was Welsh. (Heim 1982:27)

Since variable introduction (i.e., $\exists x$) brought about by the indefinite is non-deterministic, the pronoun in the second clause only refers to the non-deterministically introduced doctor value. There is no implication that London only had a doctor, who happened to be Welsh.

By contrast, a bare noun phrase must give rise to a uniqueness implication. For example, the first clause in (88), necessarily makes relevant all the children that Mingzai saw in the park (see also Cohen and Erteschik-Shir (2002) for a similar finding for English bare plurals). The plural pronoun in the second clause then refers to this maximal set of children. As a result, it is implicated that all the children were playing swing.

(88) Mingzai hai gongjuan gindou siupangjau. Keoidei wan-gan cincau.
Mingzai in park see child they play-PROG swing
'Mingzai saw children in the park. They were playing the swing.'

The uniqueness effect is not a special property of the plural pronoun. If a singular pronoun is used, a similar uniqueness implication is still observed. Consider (89). The singular pronoun gives rise to the uniqueness implication that Mingzai only saw one child and the child was playing swing.

(89) Mingzai hai gongjuan gindou siupangjau. Keoi wan-gan cincau.
Mingzai in park see child he play-PROG swing

'Mingzai saw one or more children in the park. He was playing the swing.'

The uniqueness implication associated with bare noun phrases not only indicates that the discourse status of bare noun phrases is different from that of indefinites, it also points to a plausible way to analyze the felicitous anaphora involving bare noun phrases. In particular, Chierchia (1992) observes that an indefinite in a donkey sentence is ambiguous between a strong (i.e., unique) and weak (i.e., non-unique) reading. He further proposes to distinguish between two types of anaphora. The non-unique anaphora can be derived via standard use of d-refs whereas the unique anaphora can be derived via a E-type strategy not involving the use of d-refs (see also Heim 1990). Given that anaphora involving bare noun phrases pattern like the strong reading in terms of the uniqueness implication, it is possible to extend the E-type strategy formulated for the latter to the former. However, I reserve the precise analysis for another study.

4.4.4 Indefinites with a bound pronoun

Recall from the discussion in the previous section that the presence of *saai*-distributivity and cardinal indefinites with a bound pronoun results in conflicting scopal requirements. In order to satisfy the narrow-scope distributivity requirement, the indefinite must be interpreted outside the scope of distributivity. However, in order for the pronoun to be properly bound and co-vary with distributivity, it must be inside the scope of distributivity.

Interestingly, the fact that DPlL is plural and dynamic provides a way to resolve this dilemma. Recall that the former trait allows it to represent dependencies using a plurality of assignments while the latter allows it to pass those dependencies from context to context. Now, if we let indefinites with a bound pronoun *zigei* 'self' be interpreted outside the scope of a distributivity operator but allow it to introduce its own dependency, then we can account for the fact that a pronoun-containing indefinite may co-vary with the distributivity key. More concretely, a pronoun bound by the distributivity key may induce a dependency between the distributivity key and the indefinite that contains the pronoun. Since DPlL is a dynamic logic, the dependency is passed down the stream of interpretation. When a distributivity operator is evaluated, the dependency induced by the pronoun is preserved. However, since the indefinite is introduced outside the scope of the distributivity operator, *saai* spares it for the independence constraint.

As a first step of the illustration, let us define an indefinite containing *zigei* 'self' as in (90), following van den Berg (1996)'s relational assignment $\exists y_{Rx}$.

(90) **a book^{y} zigei$_{x}$ like** $:= \lambda P.\exists y_{Rx} \wedge \mathsf{bk}\ y \wedge \mu_{\text{CARD}}\ y = 1 \wedge \mathsf{like}\ y\ x \wedge P\ y$

(91) $$G[\![\exists y_R x]\!]H := \mathbb{T} \text{ iff } H = \bigcup_{a \in G(x)} \{g^{y \to d} \mid g \in G \,\&\, R(d, a) \,\&\, d \in D_e\}$$
$$\text{iff } G(x) = H(x) \,\&\, \forall a \in G(x).G|_{x=a}[\![\exists y \wedge R\ y\ x]\!]H|_{x=a}$$

Relational assignment is formally defined in (91), which is equivalent to distributively introducing a new variable y by splitting the input info-state along the x-dimension and checking that x and y stand in a certain relation R.[17]

There are two sources of support for treating *zigei* as inducing relational assignment. First, although there has not been any study showing that the reflexive pronoun *zigei* may introduce distributivity, its close correlate in Mandarin, i.e., *ziji*, has been argued, by Huang (2002), to introduce distributivity into plural predication, based on the contrast between English reflexive pronouns and *ziji*. According to Huang (2002), plural reflexive pronouns like *themselves* in (92) is compatible with a group-praising scenario (in which every boy praised the group they belong but not himself) and a self-praising scenario (in which every boy praised himself). However, *ziji* in (93) is only compatible with the self-praising scenario. For this reason, Huang (2002) argues that that *ziji* is inherently distributive.

(92) The boys praised themselves.

(93) Nanhai-men kuajiang-le ziji.
boy-PL praise-ASP self
'Each boy praised himself.'

[17]It is more conspicuous to use a full-blown dynamic GQ translation to see that relational assignment only induces distributive evaluation in the restriction, i.e., the variable y':

(i) $\lambda P.\mathsf{max}^{y' Rx}(\mathsf{book}\ y' \wedge \mathsf{like}\ y'\ x) \wedge \mathsf{max}^{y \sqsubseteq y'}(P\ y) \wedge \mathsf{one}(y', y)$

However, the dependency is passed from y' to y because y is a dependency-preserving subset assignment of y'.

Cantonese *zigei*, when used in a plural predication, as in (94), behaves just like its Mandarin correlate *ziji* in being only compatible with the self-praising scenario.

(94) Di-Nanzai zan-zo zigei.
boy-PL praise-ASP self
'Each boy praised himself.'

Second, previous studies that use choice functions to analyze indefinites have suggested using skolemization to treat indefinites with bound pronouns (e.g, Kratzer 1998). Although I treat indefinites as dynamic generalized quantifiers rather than choice functions, relational assignment can be seen as the correlate skolemization in DPlL: a variable may be introduced to stand in a certain relation with another variable when there is explicit relational information.[18]

Combining (90) and (91), we get the definition in (95) for a cardinal indefinite with a bound pronoun:

(95) $\textbf{a book}^y\ \textbf{zigei}_x\ \textbf{like} := \lambda P.\delta_x(\exists y \wedge \mathsf{bk}\ y \wedge \mu_{\text{CARD}} y = 1 \wedge \mathsf{like}\ y\ x) \wedge P\ y$

Now, we are ready to feed this cardinal indefinite back into a sentence with a distributivity operator and *saai*. There are two positions for interpreting the cardinal indefinite, inside the scope of the distributivity operator or outside of it, as shown in the two LF configurations in Figure 4.3.

If the cardinal indefinite is interpreted *outside* the scope of distributivity (i.e., the second δ_x in (96)), *saai* ignores it for the purpose of the independence constraint. Since the indefinite has a reflexive pronoun inside it, it introduces its own distributivity (the first δ_x in (96)), scoping over the restrictor of the indefinite. This ensures that the variable introduced by the cardinal indefinite is distributively evaluated and hence may co-vary with the distributivity key.

(96) $\mathsf{max}^x(\mathsf{stds}\ x) \wedge \delta_x(\exists y \wedge \mathsf{bk}\ y \wedge \mu_{\text{CARD}} y = 1 \wedge \mathsf{like}\ y\ x) \wedge \delta_x(\mathsf{buy}\ y\ x)$

If the cardinal indefinite signaling relational assignment is interpreted *inside* the scope of distributivity, the distributivity contributed by the indefinite would be vacuous, as it falls inside the scope of

[18] The dependency-introducing random assignment defended in Brasoveanu (2007, 2008) can be seen as an unrestricted use of relational assignment.

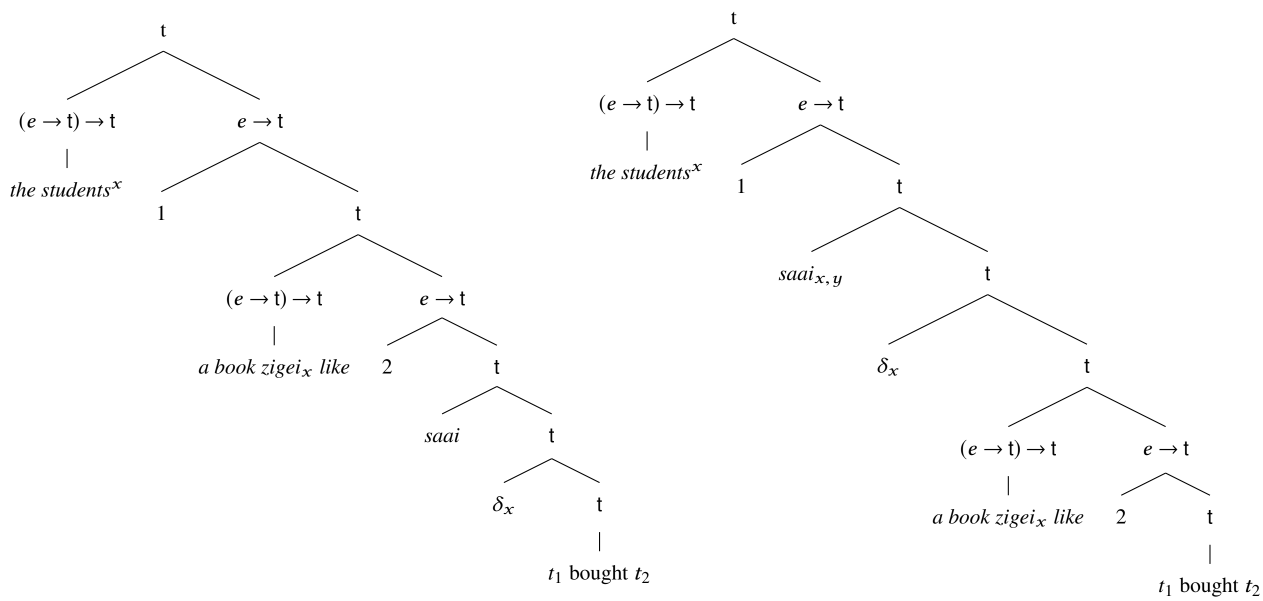


Figure 4.3: Indefinites with a bound pronoun

another distributivity operator targeting the same variable.

(97) $\mathsf{max}^x(\mathsf{stds}\ x) \wedge \delta_x(\delta_x(\exists y \wedge \mathsf{bk}\ y \wedge \mu_{\textsc{card}} y = 1 \wedge \mathsf{like}\ y\ x) \wedge \mathsf{buy}\ y\ x) \wedge \mathsf{ind}_{x,y}$

In summary, I have argued that *saai* imposes an independence constraint on the functional dependency induced by distributive quantification. I have shown how such a constraint can be couched in DPlLM, a semantics that represents the functional dependencies engendered by distributive quantification. In the next subsection, I discuss how the analysis proposed in this subsection can be compositionally implemented.

4.4.5 Compositional implementation

I assume that distributivity in a sentence with *saai* is contributed by a distributivity operator δ_x adjoined at the sentence level. The domain of the distributive quantification is determined anaphorically, by the subscripted index. This distributivity operator may be realized covertly or overtly as *cyunbou* or *dou*. It may enter into scope interactions with other noun phrases in the VP, as shown in Figure 4.4.

Saai is modeled as another sentence-level operator, located immediately above the distributivity

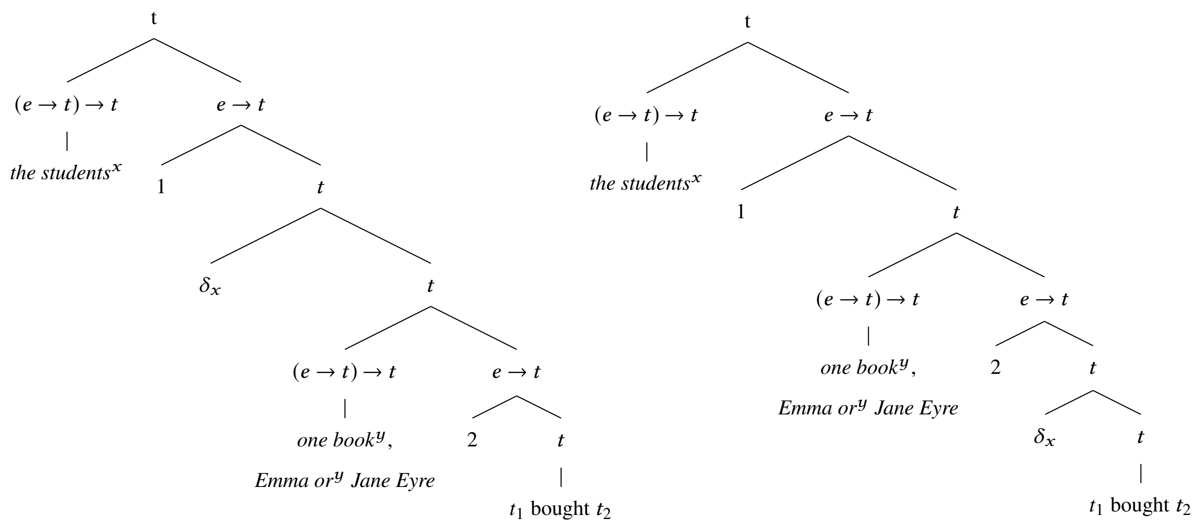


Figure 4.4: Scope interactions with distributivity

The distributivity operator may scope over (left) or under (right) a noun phrase.

operator, as shown in 4.5. It takes a dynamic proposition as its argument and imposes an independence constraint on this dynamic proposition, as shown in (98).

(98) $\textbf{saai}_{x,\vec{y}} := \lambda\phi.\phi \barwedge \textsf{ind}_{x,\vec{y}}$

Note that the constraint is imposed on a dynamic proposition with distributivity. Using ϕ as a stand-in for a distributive sentence and ψ as a stand-in for the independence constraint, we can formulate their combined contribution as follows:

(99) $G[\![\phi \barwedge \psi]\!]H = \mathbb{T}$ if $G[\![\phi]\!]H = \mathbb{T}$ & $H[\![\psi]\!]H = \mathbb{T}$, undefined otherwise.

4.5 Extension: Measurement sensitivity

4.5.1 Value independence vs. structure independence

The independence constraint is formulated in direct opposition to the well-known dependence requirement of distributive indefinites/numerals. For a distributive numeral, it is important that it does not introduce the same value relative to the distributivity key (Farkas 1997, Farkas 2002a,b Henderson 2014); however, for *saai*, it is important that a post-*saai* expression introduces the same value relative to the distributivity key.

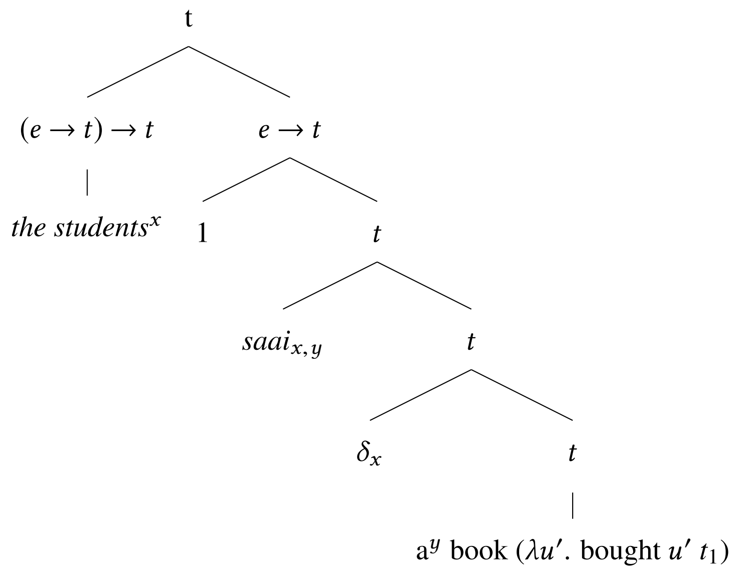


Figure 4.5: *Saai* is structurally higher than the distributivity operator

A recent finding regarding a subclass of distributive numerals, as represented by cardinal indefinites marked by binominal *each*, is their sensitivity to types of measurement (Zhang 2013, Chapter 3 of this dissertation). For example, binominal *each* requires its host to contribute an extensive measure function rather than an intensive one.

(100) The boxes are 10 pounds each.

(101) *The drinks are 90 degrees (Fahrenheit) each.

This finding has been used to argue, in Chapter 3, for an analysis in which the dependence condition of binominal *each* manifests as a monotonicity constraint checked relative to the internal mereological structure of distributivity. This type of dependence is termed **structure dependence**, to distinguish it from value dependence, which is not checked in relation to a functional dependency but not its internal mereological structure.

One may justly wonder if *saai*'s independence constraint is one of value independence or structure independence. I address this question in three steps.

- First, I show that when a discourse variable stores ordinary individuals, whether the independence constraint is stated relative to the functional dependency of distributivity or its internal, mereological structure does not make a difference.

- Second, I show that when a discourse variable stores values of degrees, then value dependence and structural dependence yield distinct predictions.
- Third, I draw on Cantonese data to show that measure phrases following *saai* are required to contribute an intensive measure function rather than an extensive one, in contrast to the requirement of binominal *each*.

The independence constraint formulated in the previous section is re-stated in (102) with the prefix V indicating that it expresses value independence. It is checked by making reference to x and y, the former allows it to associate with distributivity, and the latter allows it to target a potential y in the distributive scope of x.

(102) $G[\![\textsf{V-ind}_{x,\vec{y}}]\!]H := G = H \ \& \ \forall a,b \in G(x) : G|_{x=a}(\vec{y}) = G|_{x=b}(\vec{y})$
'y's value is constant relative to x's value.'

This constraint can be upgraded so that it is checked in association with the internal structure of distributivity. To do so, we just need to project x stored in increasingly bigger sub-info-states, and check if the corresponding values stored in y remain constant in these sub-info-states, as demonstrated in (103) (the prefix S signals that that the independence is stated in terms of structural independence).

(103) $G[\![\textsf{S-ind}_{x,\vec{y}}]\!]H := G = H \ \& \ \forall X,X' \in G(x) : G|_{x\in X}(\vec{y}) = G|_{x\in X'}(\vec{y})$
'y's value is constant relative to x's size.'

It is easy to see that (103) entails (102) when y stores individual values: if y stores the same individual(s) regardless of x's value, then y stores the same individual(s) regardless of *how many* values are assigned to x. For concreteness, let me illustrate their equivalence using the info-states in Figure 4.6.

The info-state G satisfies both value independence and structural independence: it satisfy the former because g_1 and g_2 assign different values to x but the same value to y, and it satisfies the latter

G	x	y
g_1	child1	book1
g_2	child2	book1

$$G|_{x=\mathsf{child1}}(y) = \{\mathsf{book1}\}$$

$$G|_{x=\mathsf{child2}}(y) = \{\mathsf{book1}\}$$

$$G|_{x\in\{\mathsf{chd1},\mathsf{chd2}\}}(y) = \{\mathsf{book1}\}$$

G'	x	y
g'_1	child1	book1
g'_2	child2	book2

$$G'|_{x=\mathsf{child1}}(y) = \{\mathsf{book1}\}$$

$$G'|_{x=\mathsf{child2}}(y) = \{\mathsf{book2}\}$$

$$G'|_{x\in\{\mathsf{chd1},\mathsf{chd2}\}}(y) = \{\mathsf{book1},\mathsf{book2}\}$$

Figure 4.6: G satisfies both value independence and structural independence, while G' does not satisfy either.

because when G has more assignments assigning values to x, the values G assign to y remain unchanged. The info-state G' fails to satisfy either value independence or structural independence: it fails the former because g'_1 and g'_2 do not assign the same value to y', and it fails the latter because G' assigns more values to y when there are more assignments assigning values to x. Because of this equivalence, it is impossible to tell apart value independence and structural independence if we only consider variables of individuals.

When the information stored in a discourse variable concerns **degrees** rather than individuals, then it makes a difference whether it is required to be independent at the value level or at the structure level. The reason, as we have seen in Chapter 2, is because depending on the type of measurement that produces a degree, a degree may or may not track the internal mereological structure of the individuals being measured. In the case of an extensive measurement, the corresponding degree tracks the mereological structural of the individual it measures. However, in the case of an intensive measurement, the corresponding degree does not.

Consider the two info-state in Figure 4.7, where d is a discourse variable storing degrees resulting from an extensive measurement (volume), and d' is one storing degrees from an intensive measurement (temperature (in Fahrenheit)). Degrees are modeled as triples, as discussed in Chapter 2

Intuitively, when asked how much collective volume the two drinks have and what collective temperature they have, the answer should be '12oz' and '60F' (or 'not sure'), respectively. This is because two weights can be added up straightforwardly to a bigger weight but two temperatures cannot be added up straightforwardly to a bigger temperature. As already discussed in Chapter 2,

G	x	d	d'
g_1	drink1	$\langle$6oz, vol, drink1$\rangle$	$\langle$60F, temp, drink1$\rangle$
g_2	drink2	$\langle$6oz, vol, drink2$\rangle$	$\langle$60F, temp, drink2$\rangle$

$G|_{x=\text{drink1}}(d) = \langle \text{6oz}, \text{vol}, \text{drink1} \rangle$

$G|_{x=\text{drink2}}(d) = \langle \text{6oz}, \text{vol}, \text{drink2} \rangle$

$G|_{x \in \{\text{drnk1}, \text{drnk2}\}}(d) = \langle \text{12oz}, \text{vol}, \text{dnk1} \oplus \text{dnk2} \rangle$

$G|_{x=\text{drink1}}(d) = \langle \text{60F}, \text{temp}, \text{drink1} \rangle$

$G|_{x=\text{drink2}}(d) = \langle \text{60F}, \text{temp}, \text{drink2} \rangle$

$G|_{x \in \{\text{dnk1}, \text{dnk2}\}}(d) = \langle \text{60F}, \text{temp}, \text{dnk1} \oplus \text{dnk2} \rangle$

Figure 4.7: Extensive vs. intensive measurement

this intuition can be captured by modeling degree projection not using sets, just operations on sets.

We can check that x and d in the info-state in Figure 4.6 satisfy value dependence:

(104) $\textsf{V-ind}_{x,d} := \forall a, b \in G(x) : G|^{i=1}_{x=a}(d) = G|^{i=1}_{x=b}(d)$

‘d’s value on the first coordinate is fixed/constant relative to x’s value.’

However, they do not satisfy structure independence:

(105) $\textsf{S-ind}_{x,d} := \forall X, X' \subseteq G(x) : G|^{i=1}_{x \in X}(d) = G|^{i=1}_{x \in X'}(d)$

‘d’s value on the first coordinate is fixed/constant relative to different sizes of x.’

This is because although different assignments in G assign the same value to d, collectively they assign a *different* value (a larger value, to be more precise) to d.

The situation is reversed with intensive measurement. Consider the relationship between x and d' in the info-state in Figure 4.7. They satisfy both value independence (106) and structure independence (107).

(106) $\textsf{V-ind}_{x,d} := \forall a, b \in G(x) : G|^{i=1}_{x=a}(d) = G|^{i=1}_{x=b}(d)$

‘d’s value on the first coordinate is fixed/constant relative to x’s value.’

(107) $\textsf{S-ind}_{x,d} := \forall X, X' \subseteq G(x) : G|^{i=1}_{x \in X}(d) = G|^{i=1}_{x \in X'}(d)$

‘d’s value on the first coordinate is fixed/constant relative to different value sizes of x.’

What the discussion in this subsection amounts to is that only by testing discourse variables storing degrees do we stand a chance for testing whether the independence constraint of *saai* is one of value independence or structural independence. This is because *degrees* is the sort of information that may receive the same value from every assignment but get a different value when more than one assignment is considered. An immediate question, however, is whether or not it is reasonable to assume that measure phrases introduce discourse variables over degrees. This next subsection is devoted to demonstrating that measure phrases do make dynamic contribution of degrees.

4.5.2 The dynamics of degrees

To test the dynamic contribution of measure phrases, we can test if they license donkey anaphora and cross-sentential anaphora involving degrees. Noun phrases have been argued to make dynamic contributions because they license these two types of anaphora involving individuals. The following data show that ordinary indefinites and proper names support anaphoric pronouns:

(108) Jyugwo jat-go hoksaang^x gindou jat-go lousi^y, keoi$_x$ jatding jiu heong keoi$_y$
if one-CL student see one-CL teacher, he necessarily must to him
daziufu.
greet
'If a student meets a teacher, he or she must greet him or her.'

(109) Mingzai seong gin Keongzai^x. Siufan dou seong gin *keoi*$_x$.
Mingzai want see Keongzai. Siufaan also want see him
'Mingzai wanted to see keongzai. Siufaan also wanted to see him.'

Measure phrases in Cantonese also support anaphoric reference to measurement:

(110) Jyugwo nei sik jat-bong^d jeok, nei jatding jiu sik-faan gam.do$_d$-ge coi.
if you eat one-CL meat you necessarily must eat-ALSO that.much-GE vegetables
'If you eat a pound of meat, you must eat that much vegetables.'

(111) Jyugwo nei zou jat-go zung^d wundong, nei jau jiu tau-faan gam.leoi$_d$.
if you do one-CL hour exercise you then must rest-ALSO that.long
'If you exercise for an hour, you need to rest that long.'

(112) Jyugwo nei haang saam-gonglei^d heoi hokhaau, nei jaujiu haang-FAAN gam.juan$_d$
If you walk three-kilometer to school, you need walk-ALSO that.far
faan ukkei
return home
'If you walked three kilometers to go to school, you have to walk that far to get back

home.'

(113) Mingzai sik-zo leong-go^{d} pingguo. Siufan dou sik-zo *gam.duo*$_{d}$ pingguo.
Mingzai eat-ASP two-CL apple. Siufan also eat-ASP that.much apple
'Mingzai ate two apples. Siufan also ate that many apples.'

(114) Mingzai diu-zo saam-jat^{d} jyu. Siufan dou diu-zo gam.leoi$_{d}$ jyu.
Mingzai fish-ASP three-day fish Siufan also fish-ZO that.long fish
'Mingzai fished for three days. Siufan also fished for that long.'

(115) Mingzai paau-zo saam gonglei^{d}. Siufan dou paau-zo gam.juan$_{d}$.
Mingzai run-ASP three km Sifan also run-ASP that.far
'Mingzai ran three kilometers. Sifan also ran that much.'

Given the parallelism between degrees and individuals with respect to their ability to support donkey and cross-sentential anaphora, we have reasons to believe that measure phrases and ordinary noun phrases both make dynamic contributions. Consequently, we also have reasons to suspect that the independence constraint of *saai* may interact with measure phrases. In the next subsection, I provide data from Cantonese showing that the dynamic effects of measurement indeed interact with *saai*.

4.5.3 Interactions of *saai* and measure phrases

To ease into the interaction of *saai* and measure phrases, first observe that that extensive measure phrases may occur following a verb and a verbal suffix, as shown in (116).

(116) Di-gunzong haam-zo jat-ci.
CL.PL-audience cry-ZO one-time
'The audience cried once.'

(117) Di-bengjan faan-zo leong-go jung gaau
CL.PL-patients take-ASP two-CL hour sleep
'The patients slept for two hours.'

Post-verbal extensive measure phrases may fall inside the scope of a distributivity operator, such as *cyunbou* (*dou*), without causing any issue:

(118) Di-gunzong cyunbou (dou) haam-zo jat-ci.
CL.PL-audience all DOU cry-ASP one.time
'The audience all cried once.'

(119) Di-bengjan cyunbou (dou) faan-zo leong-go jung gaau
CL.PL-patients all DOU take-ASP two-CL hour sleep
'The patients all slept for two hours.'

However, when the distributivity marker is replaced by *saai*, the sentences become ungrammatical, unless the measure phrases are removed:

(120) Di-gunzong^{x} haam-saai$_{x,d}$ (*jat-ci^{d}).
CL.PL-audience cry-SAAI one-time
'The audience all cried once.'

(121) Di-bengjan^{x} faan-saai$_{x,d}$ (*leong-go jung^{d}) gaau
CL.PL-patients take-SAAI two-CL hour sleep
'The patients all slept for two hours.'

This is unexpected if the independence constraint is stated at the value level. This is because the first coordinate of each d-value remains constant relative to increasingly more x-values in these two examples. However, once we take the independence constraint to be stated at the structural level, then the ungrammaticality falls out: the first coordinate of d indeed changes (i.e., increases) with more x values.

What about intensive measure phrases associated with an intensive measurement? Ideally, we should show that *saai* is fully compatible with intensive measure phrases and use this fact to further support that the independence constraint is stated at a structural level. However, there are a few complexities in making this argument.

First, let us first establish some cases of intensive measurement in Cantonese. The following examples show that intensive measure phrases typically occur as modifiers inside noun phrases (the relevant noun phrases are enclosed in "[...]"):

(122) Siufan maai-zo [18K-ge gaaizi].
Siufaan buy-ASP 18K-MOD ring
'Siufan bought (one or more) 18-Karat gold ring(s).'

(123) Mingzai jam-zo [sei-dou-ge binseoi].
Mingzai drink-ASP 4-DEGREE-MOD icy.water
'Mingzai drank 4-degree icy water.'

When *zo* is replaced by *saai*, the examples are indeed still fully acceptable:

(124) Di-guhaak maai-saai [18K-ge gaaizi].
CL-customers buy-SAAI 18K-MOD ring
'The customers all bought (one or more) 18-Karat ring(s).'

(125) Di-siupangjau jam-zo [sei-dou-ge binseoi].
CL-children drink-ASP 4-degree-MOD icy.water
'The children all drank 4-degree icy water.'

However, we cannot directly conclude, based on the above data, that intensive measurement supports *saai*, because the relevant noun phrases are also bare noun phrases. It is possible that they contribute kind terms and the intensive measure phrases only serve to modify the kind terms. Given that we have already seen that bare noun phrases may satisfy the independence constraint of *saai* by contributing a kind-level d-ref, we cannot be entirely sure that the degree information plays a decisive role.

To draw a more convincing conclusion, it is necessary to consider intensive measure phrases that do not serve as a modifier of a bare noun phrase. Fortunately, Cantonese has a class of measure predicates that may directly take a measure phrase as its argument. These measure predicates may take an extensive measure phrase, as in the case of *sau* 'lose, be thin', which takes an extensive measure phrase *m-bong* '5 pounds' in (126). Other measure predicates, such as *siu-dou* 'heat to' in (127), may take an intensive measure phrase, such as *100-du* '100 degrees'.

(126) Leidi wuijyun cyunbou dou *sau*-zo m-bong la.
these member all dou lose-ASP five-pounds SFP
'These members all lost five pounds.'

(127) Leidi sui cyunbou dou *siu-dou* 100-du la.
these water all dou heat-to 100-degrees SFP
'These waters have all been heated to 100 degrees.'

When *saai* attaches to these measure predicates, there is a contrast between an extensive measure phrase and an intensive measure phrase:[19]

(128) *Leidi wuijyun cyunbou dou sou-saai m-bong la.
these member all dou lose-SAAI five-pounds SFP
'These members all lost five pounds, so they can go to the next class.'

[19] Not all speakers perceive the contrast. In particular, speakers who judge cardinal indefinites following *saai* to be unacceptable even under a specific interpretation do not accept either sentence.

(129) Leidi sui cyunbou dou siu-dou-saai 100-du la.
these water all dou heat-TO-SAAI 100-degrees SFP
'These waters have all been heated to 100 degrees.'

To summarize, I have shown that *saai* can be followed by intensive measure phrases but not extensive measure phrases, indicating that it exhibits measurement-sensitivity, just like binominal *each*. For this reason, I have argued that the independence constraint of *saai* should be understood as requiring the independence of a d-ref relative to the internal mereological structure of the functional dependency of the distributivity key.

4.6 Comparison with the monotonic measurement constraint

The independence constraint of *saai* is both similar to and different from the monotonic measurement constraint of binominal *each*. The crucial feature they share lies in their reference to the internal mereological structure of a distributivity dependency. This feature explains why both distributivity markers are sensitive to type of measurement (albeit in distinct ways). As already pointed out in Chapter 2, to model measurement sensitivity it is necessary to make reference to the mereological structure of a distributivity dependency.

Although Cantonese *saai* and English binominal *each* both make reference to the mereological structure of a dependency, the constraints they impose on it differ in important ways. First, the monotonic measurement constraint of binominal *each* requires that a targeted expression (i.e., the host) tracks the size of the distributivity key whereas the independence constraint of Cantonese *saai* requires that a targeted expression (i.e., a post-*saai* noun phrase) not to track the size of the distributivity key. In addition, the monotonic measurement constraint requires access to a measure function provided by the target expression to construct degrees whereas the independence constraint has no requirement on the presence of a measure function in the target expression. Due to these differences, *saai* and binominal *each* require different types of expressions to satisfy their respective constraints. For the reader's reference, I summarize below how different types of expressions fare with *saai* and binominal *each* and where to find the relevant discussions. The summary is presented in Table 4.1 at the end of this section.

Counting quantifiers like *two books* in English and *jat-bun syu* 'one book' in Cantonese (referred to as a cardinal indefinite) can satisfy both the independence constraint (see Section 4.4.2 of this chapter) and the monotonic measurement constraint (see Section 3.4.3 of Chapter 3) but are subject to distinct interpretive requirements. Assuming that a counting quantifier contributes both an individual d-ref and a degree d-ref, the independence constraint requires them to both be structurally independent. What this means is that the individual d-ref should lack co-variation and the degree d-ref should store a set of degree names that do not depend on the size of the distributivity key. Since the degree d-ref stores the measurement information of the individual d-ref, when the individual d-ref lacks co-variation, the degree d-ref will remain constant regardless of the size of the distributivity key. This is how the two variables work together to satisfy the independence constraint. The situation is very different with the monotonic measurement constraint, which requires the individual d-ref to exhibit co-variation. Although binominal *each* is not posited to access the degree d-ref anaphorically, it accesses the same measurement information by compositionally retrieving a measure function from its host and applying the measure function to the individual d-ref provided by the host.

Measure phrases can satisfy both constraints, too, for they contribute a measure function to the monotonic measurement constraint (Section 3.4.3 of Chapter 3) and a degree variable to the independence constraint (Section 4.5 of this chapter). However, it must be made clear that it is extensive measure phrases that satisfy the monotonic measurement constraint but intensive measure phrases that satisfy the independence constraint.

Quantifiers that are hypothesized to lack an appropriate measurement component, such as *some NPs, few NPs* and *most NPs*, do not support the monotonic measurement constraint (see Section 3.3.2 of Chapter 3). However, the Cantonese correlates of these quantifiers can support the independence constraint, as long as they receive a specific interpretation (see Section 4.4.2 and footnote 2 of this chapter). In other words, they pattern like counting quantifiers with respect to the independence constraint. These quantifiers are acceptible because the independence constraint may, but need not, make use of measurement information, unlike binominal *each*.

Relatedly, disjunction such as *John or Mary* does not support the monotonic measurement constraint for its lack of a measure function. However, its correlate in Cantonese may satisfy the independence constraint as long as disjunction does not co-vary with distributivity (see Section 4.4.2 of this chapter).

Proper names such as *John* in English and *Mingzai* in Cantonese and definite expressions such

Expression	*Saai*	*Each*
Counting quantifier	✓(no co-variation)	✓(co-variation)
Extensive measure phrases	✗	✓
Intensive measure phrases	✓	✗
Non-counting quantifiers	✓(no co-variation)	✗
Disjunction	✓(no co-variation)	✗
Proper names	✓	✗
Definite expressions	✓	✗
Bare noun phrases	✓	✗

Table 4.1: Expressions (in the distributed share) that support and do not support Cantonese *saai* and English binominal *each*

as *the (two) books* in English and *go-bun syu* 'that book' in Cantonese satisfy the independence constraint (see Section 4.4.2 of this chapter) but not the monotonic measurement constraint (see Section 3.2.2 of Chapter 3). This is because these expressions do not co-vary with distributivity. The lack of co-variation allows a definite expression to readily satisfy the independence constraint. However, since co-variation is a necessary (but not sufficient) condition for the monotonic measurement constraint, the lack of co-variation makes it impossible for these expressions to satisfy the monotonic measurement constraint.

Lastly, bare noun phrases do not satisfy the monotonic measurement constraint but satisfy the independence constraint. Bare noun phrases are modeled as proper names of kinds. As such, they pattern like proper names for their inability to satisfy the monotonic measurement constraint (see Section 3.2.2 in Chapter 3 for the data) and their ability to satisfy the independence constraint (see Section 4.4.3 of this chapter).

4.7 Remaining issues

There are a few properties of *saai* that have not been addressed in this dissertation. I document them in this section to facilitate future research. To begin with, events in the scope of distributivity are not required to lack co-variation with the distributivity key. More concretely, in a sentence like (130), there is no requirement that there is only a single leaving event.

(130) Di-hoksaang zau-saai.

CL.PL-student leave-SAAI

'The students each left.'

If (130) has the LF in (131-a) and the interpretation in (131-b), it is predicted that the event d-ref introduced inside the scope of the distributivity operator should only have a single value relative to different values in the distributivity key. In other words, all the students left in the same leaving event. However, there is no requirement on what values the event d-ref may be associated with.

(131) a. $\textbf{the students}^x$ ($\textbf{cyunbou/dou}_x$ ($\lambda u.$ **leave** u)))

b. $\mathsf{max}^x(\mathsf{stdts}\ x) \wedge \delta_x(\exists e \wedge \mathsf{leave}\ e \wedge \mathsf{ag}\ e = x)\ \overline{\wedge}\ \mathsf{ind}_{x,e}$

There are a few plausible explanations. First, the constraint may be ruled out for pragmatic reasons. If the independence constraint in (131-b) took effect, it would lead to a contradiction. As pointed out in Carlson (1998) and subsequent studies, events with distinct participants are distinct events.[20] In (131-b), the events stored in *e* each has a different student as its agent. For this reason, the events cannot be independent of the size of the distributivity key—the number of events depend precisely on the number of agents found in the distributivity key. Since imposing the independence constraint on an event variable in the scope of a distributivity operator always leads to a contradiction, a pragmatic mechanism may prevent the constraint from applying to an event variable. In addition, *saai* may be only sensitive to d-refs introduced by noun phrases (including individuals and degrees). Finally, *saai*'s surface position may play a more prominent role in determining what is tracked for the independence constraint. For example, *saai* may only track the d-refs introduced by a post-*saai* expression. Since verbal predicates linearly precede *saai*, the event variables presumably introduced by them are spared.

Another interesting feature of distributivity with *saai* is that *saai* can signal distributive quantification over any grammatical position (Lee 1994, Lee 2012). For example, the distributive quantification is over the plural subject in (132) but over the plural object in (133). There is no need to

[20] I thank Robert Henderson (p.c.) for pointing this out to me.

move *saai* to a different position to indicate the change of the distributivity key.

(132) Di-hoksaang gin-saai Miss Cheung.
CL.PL-student see-SAAI Miss Cheung
'Each student saw Miss Cheung.'

(133) Miss Cheung gin-saai di-hoksaang.
Miss Cheung see-SAAI CL.PL-student
'Miss Cheung saw each student.'

Other distributivity markers, such as *cyunbou* and *dou*, are more restricted, as they can only signal distributive quantification over an expression they follow. It is possible that the independence constraint has a connection with *saai*'s flexibility with the distributivity key. I reserve this potential connection for future research.

4.8 Conclusion

In this chapter, I have used *saai* as a case study to show that there are distributivity markers that require a lack of co-variation between expressions in the distributed share and the distributivity key. Although at first glance, *saai* seems to be an entirely different beast as markers of distributive numerals, I have argued that their semantics are in fact very similar. Both types of distributivity markers impose constraints on the functional dependencies arising from distributive quantification. The only major difference is that *saai* requires a functional dependency to lack co-variation, while markers of distributive numerals require a functional dependency to exhibit co-variation.

In addition, I have shown that the independence requirement of *saai*, just like the dependence requirement of *each*, should both be understood at the structural level, rather than at the value level.

5

GENERALIZED MONOTONICITY: MANDARIN *ge* AND ADVERBIAL *each*

5.1 Introduction

This chapter is primarily devoted to the Mandarin distributivity marker *ge*. However, the conclusions drawing from the investigation of *ge* in Section 5.2 and Section 5.3 can be extended to understand determiner and adverbial *each*, which I take up in Section 5.4.

Although *ge* is widely regarded as a distributivity operator (Lin 1998b, Lee et al. 2009a), it exhibits two classes of properties that pose a challenge to this view. The first class of properties, reported in Section 5.2.1, suggests that *ge* lacks its own distributivity force, as it may co-occur with other distributivity markers that encode different types of distributivity. The second class of properties, reported in Section 5.2.2, shows that *ge*, like English binominal *each* and Cantonese suffix *saai*, imposes restrictions on what expressions can show up in a distributed share and how they are interpreted (Lin 1998b, Soh 2005, Lee et al. 2009a, Tsai 2009, Li and Law 2016). It is also shown that although these restrictions do not fully overlap with binominal *each*'s licensing conditions, they share important similarities that warrant a unified analysis.

Based on these properties, I propose (in Section 5.3) that *ge*, like binominal *each*, is a marker requiring a monotonic mapping from one structure (the mereological structure provided by the distributivity key) to another structure (the mereological structure provided by a relevant part of the distributed share). This requirement is formulated as a monotonicity constraint, which accesses the relevant mereological structures using d-refs and quantificational subordination. However, unlike binominal *each*, which is only acceptable when a monotonic mapping is from a structure consisting of individuals to a structure consisting of degrees, *ge* is compatible with more than one type of such monotonic mapping. In particular, *ge* allows mappings from individuals to degrees as well as mappings from individuals to individuals. The cross-categorial nature of *ge*'s monotonicity constraint is formally captured by allowing *ge* to use both degree d-refs and individual d-refs to retrieve a mereological structure for building the monotonicity constraint.

Lastly, in Section 5.4 I show that the monotonicity constraint can be further generalized to model the so called 'event differentiation' condition of determiner *each*, which is first discussed in detail by Tunstall (1998) (see also Vendler 1962, Brasoveanu and Dotlačil 2015). I show that by allowing a monotonicity constraint to make use of event d-refs and pragmatically available thematic functions, the event differentiation condition is just a special case of the monotonicity constraint.

5.2 The distribution of *ge*

In this section, I discuss two puzzling properties of *ge*: its co-occurrence with other distributivity markers, as well as the licensing requirements it imposes on a distributed share. These properties are collectively taken to challenge the standard view that *ge* is a distributivity marker (Lin 1998b, Lee et al. 2009a; cf. Tsai 2009).

5.2.1 Co-occurrence with other distributivity markers

Ge co-occurs with two types of distributivity markers, those marking ordinary distributivity, i.e., distributivity canonically associated with the Mandarin adverb *dou* or the English adverb *each*, and those marking 'respective' distributivity, i.e., distributivity associated with the Mandarin adverb *fenbie* 'respectively' or the English adverb *respectively*. I discuss these them in turn below.

Co-occurence with *dou*

The first piece of evidence that calls into question *ge*'s role as a distributivity operator comes from the fact that *ge* may co-occur with another distributivity marker, as noted in Tsai (2009), Lee et al.

(2009b), and Li and Law (2016). A example from Tsai (2009) is given in (1).

(1) Tamen dou ge mai-le yi-ben shu.
they DOU GE buy-ASP one-CL book
'They each bought a book.' (Tsai 2009:162)

If *dou* is a distributivity operator, as argued in Cheng (1995) and Lin (1998a), or a distributive universal quantifier, as argued in Lee (1986), it begs the question what role *ge* plays. If *ge* is also a distributivity operator contributing distributive quantification, then its distributivity should at least trigger some type of vacuous quantification effect, which is presumably responsible for the ill-formedness of English example below:

(2) *Every boy each left.

Of course, the acceptability of data points like (1) must be used with caution, as many studies have taken *dou* to not be a distributivity operator, but an operator with a semantics closer to *even* in English (Liu 2016, Xiang 2008, Xiang 2016). Given the functional multiplicity of *dou* and *ge*'s rather stable connection with distributivity, it could be argued that *ge* is the distributivity operator and *dou* is merely there to perform another function. For this reason, it is useful to look at another type of distributivity marker that can co-occur with *ge* in the next subsection.

Co-occurrence with '*respective* distributivity'

When two (or more) coordinated phrases with an equal number of conjuncts co-occur in a sentence, a special type of distributive interpretation arises, as shown in (3).

(3) Zilu he Ziyou change-le ge he tiao-le wu.
Zilu and Ziyou sing-ASP song and dance-ASP dance
'Zilu and Ziyou sang and danced, respectively.'

Since this type of distributivity can be optionally marked with the English adverb *respectively*, I refer to it as 'respective distributivity' in this study. According to the analysis advanced in Gawron and Kehler (2004) (see also Kubota and Robert 2016), this type of distributivity involves a covert distributivity operator RESP_f (I will return to this operator in Section 5.3.4). This operator takes two pluralities at a time, break them into parts, pairs the parts using a pragmatically available sequencing function f, and performs a pair-wise evaluation facilitated by f.

Although RESP_f is a covert operator, there are lexical items such as English *respectively* that can be added to force respective distributivity. In Mandarin, the adverb *fenbie* 'separately, respectively' can be used for a purpose. When there is only one coordinated phrase, it is interpreted as *separately* (see Lasersohn 1995, 1998 for English adverb *alternatively*, which has a similar interpretation):

(4) Zilu he Ziyou fenbie chang-le ge.
Zilu and Ziyou separately sing-ASP song
'Zilu and Ziyou sang a song separately.'

(5) Zilu fenbie chang-le ge he tiao-le wu.
Zilu separately sing-ASP song and jump-ASP dance
'Zilu sang and danced separately.'

When there is more than one coordinated phrase, *fenbie* serves to mark the respective distribution, as shown in (6).[1]

(6) Zilu he Ziyou fenbie change-le ge he tiao-le wu.
Zilu and Ziyou respectively sing-ASP song and dance-ASP dance
'Zilu and Ziyou sang and danced, respectively.'

Since the respective distributivity operator RESP_f is incompatible with the ordinary distributivity operator contributed by *dou*, the respective interpretation vanishes when *dou* co-occurs with *fenbie*.[2] As shown in (7), when the two co-occur *fenbie* can only take up the interpretation of *separately*.

(7) Zilu he Ziyou (dou) fenbie (dou) change-le ge he tiao-le wu
Zilu and Ziyou DOU separately DOU sing-ASP song and dance-ASP dance
'Each of Zilu and Ziyou sang and danced separately.'

However, when *ge* and *fenbie* co-occur, as shown in (8), the respective distributivity interpretation is still available. In other words, *ge*, unlike *dou*, does not introduce a distributivity operator that would

[1] The *separately* interpretation is still available but less preferred.

[2] The incompatibility is due to the fact that *dou* only breaks down one coordinated phrase, i.e., the coordinated subject, for establishing distributive quantification. RESP_f, which falls inside the scope of *dou*'s distributive quantification, only has access to one plurality, i.e., the plurality contributed by the coordinated VP. The first plurality is no longer available since it is inside the scope of *dou*. Since the respective distributivity interpretation must be built with two pluralities, it is hence not available inside the scope of *dou*.

interfere with the formation of respective distributivity.

(8)	Zilu	he	Ziyou	(ge)	fenbie	(ge)	change-le	ge	he	tiao-le	wu.
	Zilu	and	Ziyou	GE	respectively	GE	sing-ASP	song	and	dance-ASP	dance

‘Zilu and Ziyou sang and danced, respectively.’

There is more than one way to interpret the behavior of *ge*. We may take *ge* to contribute a distributivity operator, as does in Lin (1998b) and Lee et al. (2009a), and devise a mechanism to deactivate the operator when another one is present. Alternatively, we may take *ge* to embody the respective distributivity operator RESP_f and reduce all distributivity with *ge* to respective distributivity, a line of research explored in Tsai (2009). Lastly, we may take *ge* to not contribute a distributivity operator at all. On this view, it is compatible with different types of distributivity because it does not contribute a distributivity operator of its own. However, for the last view to have any traction, it is necessary to clarify a few questions: if it *ge* does not contribute a distributivity operator, why does it always show up in a distributivity sentence? What functions does it serve in a distributively interpreted sentence?

To answer these important questions, I turn to another set of *ge*'s distributional properties in the next subsection. These properties show that *ge* is not compatible with just any sentence with a distributive interpretation. In particular, *ge*'s presence needs to be licensed by certain morphosyntactic and interpretive properties of expressions in the distributed share.

5.2.2 *Ge*'s licensing requirements

In this section, I discuss *ge*'s licensing requirements. The term ‘licensing’ is borrowed pre-theoretically from the literature on negative polarity items to describe the fact that *ge* is only felicitously used when certain factors are met. I discuss the range of conditions that licenses *ge* and the generalizations we can draw from these licensing conditions. Many of the licensing conditions have been reported in the literature, in Kung (1993), Lin (1998b), Soh (2005), Lee et al. (2009a), Tsai (2009), and Li and Law (2016). Moreover, based on these licensing conditions, Lin (1998b), Lee et al. (2009a), Tsai (2009), and Li and Law (2016) have developed analyses of *ge* that are closely related to the generalized monotonicity constraint proposed in this study. I will offer a review of these studies in Section 5.5, after developing my own analysis in Section 5.3.

A point of clarification. To highlight *ge*'s licensing condition, I use an unmarked distributivity marker, i.e., *dou*, as a comparison. For the purpose of this study, I follow Lee (1986), Cheng (1995), and Lin (1998a) in treating *dou* as a distributivity marker. However, I do not rule out the possibility that *dou* is merely compatible with distributivity, rather than contributing distributivity (Xiang 2008, Liu 2016, Xiang 2016). If *dou* turns out to not be a distributivity marker, the differences between *ge* and *dou* will be attributed to the differences between *ge* and whatever mechanism gives rise to a distributive interpretation, such as the use of a null distributivity operator.

Licensing by counting quantifiers and measure phrases

The first category of expressions that licenses *ge* is **counting quantifiers** (Kung 1993, Lin 1998b, Tsai 2009, Lee et al. 2009a, Li and Law 2016). Observe that while the counting quantifier is obligatory when distributivity is marked by *ge*, as shown in (9), it is optional when distributivity is marked by *dou*.

(9) Zhe-xie haizi ge kan-le *(liang-chu) dianyin.
this-CL.PL child GE see-ASP two-CL movie
'The children saw two movies each.'

(10) Zhe-xie haizi dou kan-le (liang-chu) dianyin.
these-CL child DOU see-ASP two-CL movie
'The children all saw two movies.'

Recall, from Chapter 3, that counting quantifiers are also required to license distributivity with binominal *each*:

(11) The girls saw *(two) movies each.

The parallelism between binominal *each* and *ge* regarding licensing by counting quantifiers goes beyond the morphosyntactic requirement of a counting quantifier in the distributed share. In fact, they also share an important interpretive property, namely, that the counting quantifier must receive a narrow-scope interpretation and co-varies with the distributivity key. For this reason, the wide scope, specific interpretations of the counting quantifiers are unacceptable:

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

(13) ??Zhe-xie haizi ge kan-le liang-chu dianyin, jiushi Afanda he Bingheshiji
these-CL child GE see-ASP two-CL movie, namely Avatar and Ice.Age
'The children saw two movies each, namely, Avatar and Ice Age.'

Closely related to counting quantifiers are **measure phrases**. Recall that measure phrases with an extensive measure function can license the use of binominal *each* but those with an intensive measure function cannot. This contrast is shown in (14-a) and (14-b). The measure phrase in the former provides an extensive measure function, i.e., **volume** (in ounce), but the measure phrase in the latter provides an intensive measure function, i.e., **temperature**.

(14) a. The drinks are six ounces each.
b. ??The drinks are sixty degrees each.

The same contrast holds for *ge*. The extensive measure phrase with the measure function **volume** (in mimiliter) in (15-a) licenses *ge* but the intensive measure phrase with the measure function **temperature** in (15-b) does not.

(15) a. Zhe-xie sui ge (you) 200 haosheng.
these-CL water GE have 200 mililiter
'The waters are each 200 ml.'

b. ??Zhe-xie sui ge (you) 60 du.
these-CL water GE have 60 degree
'The waters are each 60 degrees.'

Considering the data presented so far, it may seem to the reader that *ge* is merely a Chinese variant of binominal *each*. It can be licensed by expressions like counting quantifiers and extensive measure phrases because they have a measure function component that *ge* needs in order to construct a monotonicity constraint. However, simply equating *ge* with binominal *each* is premature, as *ge* differs from binominal *each* in two important respects.

First, *ge* needs not be adjacent to the expression that licenses it. (16) shows that *ge* can be separated from the counting quantifier that licenses it by a main verb (and an aspectual suffix). (17) and (18) show that there can also be additional adverbials between *ge* and the counting quantifier.

(16) Zhe-xie haizi ge kan-le liang-chu dianyin.
these-CL child GE see-ASP two-CL movie
'The children each saw two movies.'

(17) Zhe-xie haizi ge *zai xinqitian* kan-le liang-chu dianyin.
these-CL child GE on Sunday see-ASP two-CL movie
'The children each saw two movies on Sunday.'

(18) Zhe-xie haizi ge *zai xinqitian toutou-de* kan-le liang-chu dianyin.
these-CL child GE on Sunday sneakily see-ASP two-CL movie
'The children each saw two movies on Sunday sneakily.'

The lack of an adjacency requirement stands in contrast to the distribution of binominal *each*, which, according to Stowell (2013), has a strong preference to immediately follow the counting quantifier that hosts it. The examples below serve to show the adjacency requirement of binominal *each*.

(19) a. The boys carefully read one book each. (Stowell 2013: (56c))
b. %The boys read one book carefully each. (Stowell 2013: (56e))

An implication of the lack of an adjacency requirement for *ge* lies in compositionality. If we are to explain *ge*'s sensitivity towards measurement type along similar lines as the monotonic measurement constraint of binominal *each*, it is necessary to find a way to extract the measure function from the measure phrase. In the case of binominal *each*, the extraction is done syntactically. This is possible because *each* forms an immediate constituent with its licensor (Safir and Stowell 1988). Since *ge* does not form an immediate constituent with its licensor, we need to find ways for *ge* to gain access to a measure function, which I assume is inside a noun phrase. There are at least two hypotheses that we can entertain. To begin with, *ge* may be underlyingly more similar to binominal *each* in being adjacent to its licensors for interpretive purposes.[3] However, it may undergoes movement to the boundary of a verb phrase for syntactic reasons. Although such an analysis has not been proposed in the literature, previous studies have argued that *ge* may adjoin to different verb phrases when there is more than one verb phrase available (Lin 1998b, Soh 2005). If the movement analysis happens to be correct, then *ge* can access a measure function syntactically, in the same way that binominal *each* accesses one. Alternatively, *ge* may extract the measure function at a distance

[3] I thank Veneeta Dayal (p.c.) for pointing out this possibility to me.

using discourse anaphora. In Section 5.3.1, I propose an analysis in terms of discourse anaphora involving a dependent degree variable to achieve this effect.

The second difference between *ge* and binominal *each* lies in their licensing conditions. While binominal *each* is only licensed by counting quantifiers and extensive measure phrases, *ge* admits a wider range of licensors. In addition to counting quantifiers and measure phrases, it can be licensed if the distributed share contains any of the following expressions: a pronoun bound by the distributivity key, a quantifier-internal adjective like *butong* 'different', and an interrogative *wh*-expression inducing a pair-list interpretation. I discuss these licensing conditions in turn below.

Licensing by bound pronouns

We have seen, in (9), that a bare noun phrase does not license *ge*. The example is repeated below without the numeral or the classifier, i.e., without the counting component.[4] Since a bare noun phrase in Mandarin is ambiguous between an existential, indefinite interpretation as well as a definite interpretation (Cheng and Sybesma 1999, Yang 2001, Dayal 2013, Jenks 2018), it is necessary to clarify that neither interpretation licenses the use of *ge*.

(20) ??Zhe-xie haizi ge kan-le shu.
this-CL.PL child GE see-ASP shu
'*The children read books each.'

Adding a pronoun (but not a proper name) in the distributed share significantly improves the above sentence, as shown in (21) and (22) (see also Lee et al. 2009a, Tsai 2009).

(21) Zhe-xie haizi ge kan-le ziji/*Zilu dailiang-de shu.
these-CL.PL child GE read-ASP self/Zilu bring-DE book
'These children each read the book(s) they/??Zilu brought.'

(22) Zhe-xie haizi ge kan-le ziji-de/*Jinyong-de shu.
these-CL.PL child GE read-ASP self-POSS/Jinyong-POSS book
'These children each read their/Jinyong's book.'

[4] According to Tsai (2009), sentences like (20) become acceptable if the aspectual marker is changed to *guo*, which marks a perfective aspect of a repeatable event. The speakers I have consulted do not find the aspect marker *guo* a useful aid. A potential factor for the different judgments lies in the type of Mandarin being studied in Tsai (2009) and in the present study. While the present study is based on Mandarin as spoken in Mainland China, Tsai (2009) is likely based on Mandarin spoke in Taiwan.

The form of the pronoun is relatively flexible. It may be in the form of a reflexive pronoun *ziji*, as in (21) and (22), a third person plural pronoun *tamen* 'they', as in (23), or a reciprocal pronoun *duifang* 'the other/each other', as in (24).[5]

(23) Zhe-xie haizi ge kan-le tamen-de shu.
these-CL.PL child GE read-ASP they-DE book
'These children each read their book.'

(24) Zilu he Ziyou ge kan-le duifang-de shu.
Zilu and Ziyou GE read-ASP the.other-DE book
'Zilu and Ziyou each read the other's book.'

However, the *interpretation* of the pronoun is subject to restrictions. In particular, the pronoun must co-vary with the distributivity key. I call this observation the **bound pronoun generalization**. This generalization can be most readily verified when the pronoun involved is a third person plural pronoun, which is ambiguous between a so-called 'free' interpretation (referring to a contextually salient individual) and a so called 'bound' interpretation (co-varying with a quantifier in the same sentence). As shown in (25), when *dou* is used to give rise to distributivity, the pronoun in the distributed share may receive a bound interpretation or a free one. However, when *ge* is used to mark distributivity, as in (26), the pronoun may only receive a bound interpretation.

(25) Zhe-xie haizi^x dou kan-le tamen$_{x/y}$-de shu.
these-CL.PL child DOU read-ASP they-DE book
'These children^x each read their$_{x/y}$ book(s).' Bound/Free

(26) Zhe-xie haizi^x ge kan-le tamen$_{x/*y}$-de shu.
these-CL.PL child GE read-ASP they-DE book
'These children each read their$_{x/*y}$ book(s).' Bound/*Free

With some care, it is possible to verify the bound pronoun generalization by using a reciprocal pronoun or a reflexive pronoun. The reciprocal pronoun *duifang* is also ambiguous between a free interpretation (27) and a bound one (28), when distributivity is marked by *dou*. However, when distributivity is marked by *ge*, only the bound interpretation survives.

[5] Similar to English, a third person singular pronoun in Mandarin cannot take a morphologically plural noun phrase as its antecedent, even when the pronoun is in the scope of a distributive operator.

(27) Zilu he^x Ziyou dou kandao-le duifang$_{x/y}$-de lian.
Zilu and Ziyou DOU see-ASP the.other-DE face
'Zilu and^x Ziyou saw each other$_x$'s face/the other$_y$ person's face.' Bound/Free

(28) Zilu he^x Ziyou ge kandao-le duifang$_{x/*y}$-de lian.
Zilu and Ziyou GE see-ASP the.other-DE face
'Zilu and^x Ziyou saw each other$_x$'s face/*the other$_y$ person's face.' Bound/*Free

The reflexive pronoun *ziji* must be bound by an antecedent introduced within the same sentence (Huang 1982, Tang 1989, Pan 1998, a.o.), so a 'free' interpretation is independently unavailable. However, since *ziji* may take on different antecedents that occur before it (see also Huang 1982, Huang and Liu 2001, a.o.), the flexibility can be used to corroborate the bound pronoun generalization. Observe first that the antecedent of *ziji* is ambiguous in (29) when distributivity is marked with *dou*: *ziji* may refer to the higher subject *Zhang Laoshi* 'Teacher Zhang' or the lower subject *zhe-xie haizi* 'these children', which is also the distributivity key.

(29) Zhang Laoshi^x rang zhe-xie haizi^y dou kan-le ziji$_{x/y}$ dailai-de shu.
Zhang Teacher ask these-CL.PL child DOU read-ASP self bring-DE book
'Teacher Zhang^x asked these children^y to all read the book he$_x$/they$_y$ brought.'

However, once *ge* is used in place of *dou*, the ambiguity is gone. In (30), *ziji* can only refer to the distributivity key *zhe-xie haizi* 'these children'.

(30) Zhang Laoshi^x rang zhe-xie haizi^y ge kan-le ziji$_{*x/y}$-de shu.
Zhang Teacher ask this-CL.PL child GE read-ASP self-DE book
'Teacher Zhang asked these students to read their/her own books.'

Licensing by internal readings

In addition to bound pronouns, *ge* can be licensed by expressions in the distributed share that induces a so-called sentence-internal interpretation. I call this the **internal reading generalization**. An example is given in (31), which shows that the adjective *butong* 'different' licenses *ge*. Similar expressions like *buyiyang* 'different' have the same licensing effect.

(31) Zhe-xie haizi ge kan-le butong-de shu.
this-CL.PL child GE read different-DE book
'These children read different books.'

Carlson (1987) points out that English *different* has a sentence-internal interpretation (also known as a bound interpretation) as well as a sentence-external interpretation (also known as a free interpretation) (see also Beck 2000, Brasoveanu 2011). These interpretations are similar to the bound and free interpretations of reciprocal pronouns discussed earlier.

In a sentence with distributivity marked by *dou*, such as (32-b), *butong-de shu* 'different book(s)' may refer to a single book that is different from a salient sentence-external antecedent, i.e., *Emma* introduced in (32-a). This is the external interpretation of indicated in (32-b-i). Alternatively, it may refer to a set of different books that different children read, as shown in (32-b-ii). This is the internals interpretation.

(32) a. Laoshi kan-le *Emma*.
teacher read-ASP Emma
'The teacher read Emma'

b. Zhe-xie haizi dou kan-le butong-de shu
this-CL.PL child DOU read-ASP different-DE book
(i) The children each read some book(s) that differed from *Emma*. External
(ii) The children each read some book(s) that differed from the books the rest of the children read. Internal

When *dou* is replaced by *ge*, however, the free interpretation becomes unavailable:

(33) a. Laoshi kan-le *Emma*.
teacher read-ASP Emma
'The teacher read Emma'

b. Zhe-xie haizi ge kan-le butong-de shu
this-CL.PL child GE read-ASP different-DE book
(i) The children each read some book(s) that differed from *Emma*. *External
(ii) The children each read some book(s) that differed from the books the rest of the children read. Internal

As will be shown in Section 5.3.3, the internal reading generalization and the bound pronoun generalization are similar in nature. They both involve a variable, when evaluated inside the scope of distributivity, gives rise to co-variation.

Licensing by *respective* distributivity

As I have already discussed in Section (2), *ge* is licensed when two conjunctions give rise to respective distributivity. An example involving respective distributivity is given below:

(34) Zilu he Ziyou ge chang-le ge he tiao-le wu.
Zilu and Ziyou GE sing-ASP song and jump-ASP dance
'Zilu and Ziyou sang a song and performed a dance, respectively.'

There is evidence showing that it is the respective interpretation that licenses *ge*, but not just the use of two plural noun phrases. (35) shows that when the two conjunctions are replaced by two definite plurals, *ge* becomes highly marked.

(35) ??Zhe-xie haizi ge kan-le na-xie shu.
this-CL.PL child GE read-ASP that-CL.PL book
Intended: 'These children read those books, respectively.'

Pragmatically providing a pairing between the children and the books do not provide much help to (35). To make a respective interpretation fully acceptable with two definite plurals, the respective number of the entities contributed by the plurals have to be specified, and preferably the adverb *fenbie* is also used. When these ingredients are present, as in (36), the use of *ge* is licensed.

(36) Zhe san-ge haizi fenbie ge kan-le na san-ben shu.
this three-CL child respectively GE read-ASP that three-CL book
'These three children read those three books, respectively.'

Since the concern of the present study is on licensing *ge* rather than on licensing respective distributivity, I do not go into details as to why there is a difference between (35) and (36) in supporting respective distributivity. I the generalization to be that whenever respective distributivity is available, *ge* is licensed.

Licensing by pair-list interpretations

The last category of expressions that license *ge* is *wh*-questions with a pair-list interpretation. I would like to note that I will not develop an analysis for the pair-list interpretations of *wh*-questions involving distributive quantifiers, for modeling questions adds considerable complexity to the DPlLM framework I have been using. For readers interested in modeling the semantics of pair-list interpretations, please refer to Chierchia (1993), Krifka (2001), Dayal (1996, 2017). However, it is still useful to examine licensing by pair-list interpretations as this is another type of licensing condition that *ge* and binominal *each* differ.

Consider (37-a) first, in which distributivity is marked with *dou*. The question admits a pair-list answer, such as (37-b), as well as a single answer, such as (37-c).

(37) a. Zhe-xie haizi dou kan-le shenme shu?
this-CL.PL child DOU read-ASP what book
'What book did these children each read?'

b. Zilu kan-le Emma, Ziyou kan-le Jane Eyre, Mali kan-le Pride and Prejudice.
Zilu read-ASP Emma Ziyou read-ASP Jane Eyre Mary read-ASP Pride and Prejudice
'Zilu read Emma, Ziyou Jane Eyre, and Mary Pride and Prejudice.' Pair-list

c. Tamen kan-le Emma.
they read-ASP Emma
'They read Emma.' Single answer

However, once the distributivity marker becomes *ge*, as in (38-a), only the pair-list answer (38-b) is acceptable. The single answer (37-c) is infelicitous.

(38) a. Zhe-xie haizi ge kan-le shenme shu?
this-CL.PL child GE read-ASP what book
'What book did these children each read?'

b. Zilu kan-le Emma, Ziyou kan-le Jane Eyre, Mali kan-le Pride and Prejudice.
Zilu read-ASP Emma Ziyou read-ASP Jane Eyre Mary read-ASP Pride and Prejudice
'Zilu read Emma, Ziyou Jane Eyre, and Mary Pride and Prejudice.' Pair-list

c. #Tamen kan-le Emma.
they read-ASP Emma

'They read Emma.' #Single answer

While binominal *each* can be licensed by *wh*-expressions, it is only *how many*-NPs, which have a measurement component due to the presence of *many* (Hackl 2000, Kennedy 2015, a.o.o), that license it:

(39) How many books each did the girls read?

In fact, it is the measurement in (39) that licenses binominal *each*, rather than the pair-list interpretation. This is because a felicitous answer to this question may be a pair-list answer (e.g., *five, three, and six*) or a single answer (e.g., *five*). Using a *wh*-phrase that lacks a measurement component does not license binominal *each*, as shown in (40).

(40) *What books/which book each did the girls read?

Many studies have shown that the pair-list interpretation of a *wh*-question involving a quantifier is intimately tied to distributivity (Dayal 1996).

Interim summary

In summary, *ge* is licensed by counting quantifiers, extensive measure phrases, bound pronouns, and sentence-internal readings, and pair-list readings. A challenge in front of us is how to make sense of this conglomerate of licensing conditions. Since *ge* patterns like binominal *each* with regard to licensing by counting quantifiers and extensive measure phrases, it is reasonable to assume that *ge* also bears some form of monotonic measurement constraint. However, since *ge* can also be licensed without the presence of any measurement, the constraint must be more general than the monotonic measurement constraint.

In the next section, I develop a generalized monotonicity constraint on the basis of the monotonic measurement constraint. The idea is that *ge* is similar to binominal *each* in requiring a monotonic mapping between two pluralities living in a distributivity dependency. However, while binominal *each* only allows a mapping from individuals to measurements of individuals, i.e., degrees, to satisfy the monotonic measurement constraint, *ge* also allows a mapping from individuals to individuals to satisfy its monotonicity constraint. Because of the generality of *ge*'s constraint, I call it a generalized monotonicity constraint. I discuss this constraint in more detail in the next section.

5.3 Proposal: a generalized monotonicity constraint in DPlLM

At the heart of a monotonicity constraint on distributivity is a monotonic mapping between two mereological structures: the mereological structure contributed by a distributivity key and the mereological structure contributed by a dependent expression *after* it has been distributively evaluated. We have seen, from Chapter 1.3.1, that the monotonic measurement constraint of binominal *each* requires a monotonic mapping from the distributivity key to a dependent measurement, either contributed by a measure phrase or a counting quantifier.

(41) The drinks are 6oz each.

(42) The students bought three books each.

Using (41), the mapping can be illustrated with help of Figure 5.1.[6]

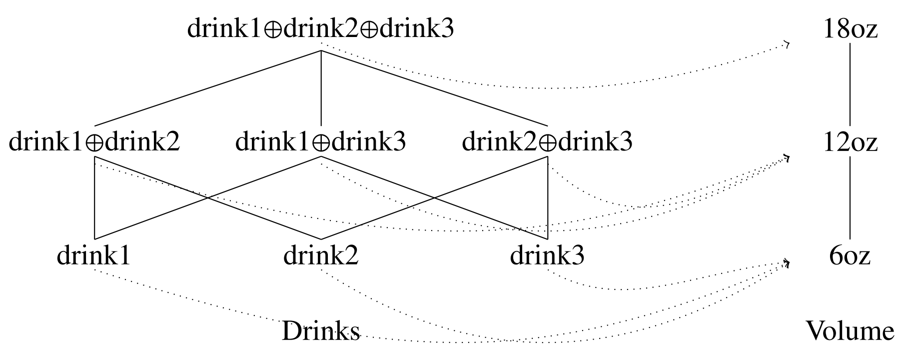


Figure 5.1: Monotonic mapping from individuals to degrees

We have also seen that quantifiers that do not bear a measure function component, such as *some NPs*, *most NPs*, and *few NPs*, do not support binominal *each*. An example is given below:

(43) *Every boy bought some books/a certain book each.

[6]Since degrees are modeled as triples in this study, this mapping is actually from individuals to the first coordinates of a set of degrees. The first coordinate of each degree stores a degree name, which is a point on a fully ordered scale.

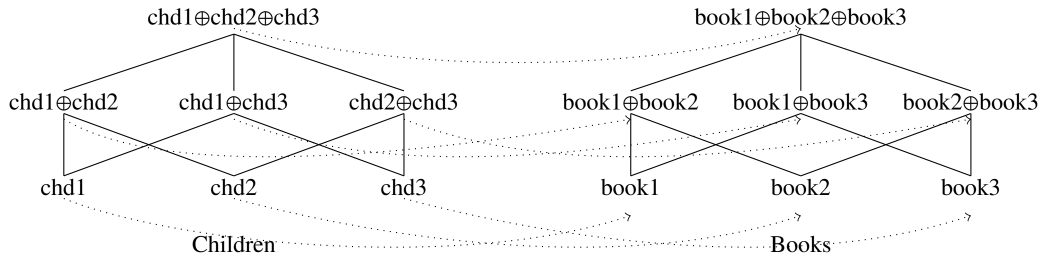


Figure 5.2: Monotonic mapping from individuals to individuals

The reason is because while these quantifiers contribute a mereological structure based on individuals, their lack of measure function does not support the building of a degree mereology. In other words, binominal *each* is selective about the type of monotonic mapping between two mereological structures: while it can be an individual-degree mapping, it cannot be an individual-individual mapping. An example of an individual-individual mapping is illustrated in Figure 5.2.

Mandarin *ge* is less selective than binominal *each*. It is compatible with both types of monotonic mappings: an individual-degree mapping, as well as an individual-individual mapping. I argue that recognizing these two types of mappings is all we need to account for the licensing conditions of *ge*.

Translating *ge*'s generalized monotonicity constraint into the DPlLM framework is quite transparent. We just need to make reference to a pair of d-refs. The first d-ref stores values contributed by the distributivity key. The second d-ref stores values contributed by a dependent expression. The constraint, as given in (44), then requires that the mereological structures computable from the values stored in the two variables observe monotonicity. The flexibility in the type of the second d-ref u is the formal reflex of the generality of the monotonicity constraint.

(44) $\mathsf{dm}_{x,u}$, where u may be a degree variable or an individual variable.

Let us first consider the case when u is resolved to a degree variable d. I have argued, in Chapters 2 and 3, that a monotonic mapping from a set individuals to a set of degrees should be understood as structural dependence, whose definition is given below:

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

a. there are distinct nonempty sets $A, B \subseteq G(x)$: $G|^{i=1}_{x\in A}(d) \neq G|^{i=1}_{x\in B}(d)$

b. for all distinct nonempty sets $A, B \subseteq G(x)$: if $A \subseteq B$, then $G|^{i=1}_{x \in A}(d) \leq G|^{i=1}_{x \in B}(d)$.

Recall that a degree is modeled as a triple. The first coordinate of the triple stores a degree name. A degree name can be retrieved with help of a parameterized projection function $G|^{i=1}(d)$, where i is the coordinate to be projected. The second coordinate of a degree stores a measure function (retrieved using $G|^{i=2}(d)$), and the third coordinate stores the individual being measured (retrieved using $G|^{i=3}(d)$). The complex structure is motivated by the need to concatenate degrees to build a degree mereology. A degree name is always derived by applying a measure function to an individual. Without the accompanying information about the measure function and the individual being measured, it is very hard to determine how two degrees are to be concatenated only by looking at the degree names.[7]

Concretely, to compute the first coordinate of a degree, i.e., $G^{i=1}(d)$, which is a degree name, we take the measure function stored in the second coordinate ($\bigoplus G^{i=2}(d)$) and apply it to be (possibly plural) individual stored in third coordinate ($\bigoplus G^{i=3}(d)$). The summation operator $\bigoplus$ is used for different reasons here. Since the second coordinate of a degree stores the same measure function for all assignments, the $G^{i=2}(d)$ is a singleton set containing one measure function. The summation operator simply removes the set and returns the measure function. However, $G^{i=3}(d)$ returns a set of individuals. The summation operator sums together all of these individuals in the set, allowing a measure function to apply to the sum individual. The following notation is used to compute a degree name from a degree d-ref in a plural info-state.

(46) $G^{i=1}(d) = \bigoplus G|^{i=2}(d)(\bigoplus G^{i=3}(d))$

The constraint in (45) differs from the monotonic measurement constraint of binominal *each* only in compositionality. Specifically, binominal *each* assembles a monotonic measurement constraint $\mathsf{dm}_{x,y}(\mu)$ by gaining syntactic access to the measure function μ, then applying the measure function to the dependent individual variable to yield a set of degree information (i.e., $\mu(y)$ for each assignment g in G). The monotonicity requirement ultimates holds between a mereological structure of

[7]This is regardless of whether we allow an intensive measure function to apply to two non-overlapping objects or not. Even if we disallow an intensive measure function to generate a defined measurement for non-overlapping objects, we still need access to such a measure function to know that the effect of concatenating two intensive measurements is undefined.

individuals and a mereological structure of degrees. To satisfy the monotonicity requirement of binominal *each*, the same conditions in (45) have to be satisfied.

If u is resolved to an individual variable, such as y, then the monotonicity constraint requires structural dependence between the distributivity key variable and the dependent individual variable.

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

a. there are distinct nonempty sets $A, B \subseteq G(x)$: $G|_{x\in A}(y) \neq G|_{x\in B}(y)$

b. for all distinct nonempty sets $A, B \subseteq G(x)$: if $A \subseteq B$, then $G|_{x\in A}\ y \subseteq G|_{x\in B}\ y$.

Recall from Chapter 2 that when a monotonic mapping is from individuals to individuals, then it can be recast in terms of value dependence or co-variation.[8] However, I maintain an analysis in terms of monotonicity to reflect that *ge* is still sensitive to the extensive-intensive distinction of measurement. The fact that when a monotonic mapping only involves individuals it can be reduced to value dependence or co-variation follows straightforwardly from the definition of the sum operation on a set of individuals: two distinct individuals can be summed to form a straightly bigger individual, while the summation of two identical individuals does not give rise to a straightly bigger individual, as it always returns the same individual.

In the next section, I discuss how the generalized monotonicity constraint of *ge* accounts for its distribution.

5.3.1 Accounting for licensing by counting quantifiers and measure phrases

Ge is licensed by measure phrases when they provide an extensive measure function, as shown in (48-a) and (48-b).

(48) a. Zhe-xie sui^x ge (you) 200 haosheng^d.
these-CL water GE have 200 mililiter
'The waters are each 200 ml.' **Volume**

b. *Zhe-xie sui^x ge (you) 60 du^d.
these-CL water GE have 60 degree
'The waters are each 60 degrees.' **Temperature**

[8] Evaluation plurality, first proposed in Henderson (2014), and later adopted in Champollion (2015) and Kuhn (2017), can be reducible to value dependence.

The dependent expressions in these examples are measure phrases. I assume that the measure phrases *(you) 200 haosheng* 'be 200ml' and *(you) 60 du* 'be 60C' are used predicatively, so they are functions from individuals to dynamic propositions. Their dynamic contribution is the introduction of a degree d-ref, which is linked to an individual by a measure function μ.

(49) **(be) 200ml** $:= \lambda x.\exists d \ \wedge \ d = \langle \mathsf{200ml}, \mathsf{vol}, x \rangle \ \wedge \ \mu\, x = d$

(50) **(be) 60C** $:= \lambda x.\exists d \ \wedge \ d = \langle \mathsf{60C}, \mathsf{temp}, x \rangle \ \wedge \ \mu\, x = d$

The plural demonstrative phrase *zhe-xie siu* 'these waters' in (48-a) and (48-b) is treated as a dynamic generalized quantifier:

(51) $\lambda P.\mathsf{max}^x(\mathsf{water}\ x) \ \wedge P\ x$

Distributivity is introduced by a covert distributivity operator:

(52) **Dist** $:= \lambda P \lambda x.\delta_x(P\ x)$

Combining the generalized quantifier, the distributivity operator, and the predicative measure phrases in (49) and (50) in the manner in (53-a) and (54-a) yields (53-b) and (54-b), respectively:

(53) a. **these waters** (**Dist** $(\lambda y.\exists d \ \wedge \ d = \langle \mathsf{200ml}, \mathsf{vol}, x \rangle \ \wedge \mu\ y = d)$)

b. $\mathsf{max}^x(\mathsf{water}\ x) \ \wedge \ \delta_x(\exists d \wedge \ d = \langle \mathsf{200ml}, \mathsf{vol}, \mathsf{x} \rangle \ \wedge \mu\ x = d) \ \wedge \mathsf{dm}_{x,d}$

(54) a. **these waters** (**Dist** $(\lambda y.\exists d \ \wedge \ d = \langle \mathsf{60C}, \mathsf{temp}, x \rangle \ \wedge \mu\ y = d)$)

b. $\mathsf{max}^x(\mathsf{water}\ x) \ \wedge \ \delta_x(\exists d \wedge \ d = \langle \mathsf{60C}, \mathsf{temp}, \mathsf{x} \rangle \ \wedge \mu\ x = d) \ \wedge \mathsf{dm}_{x,d}$

As already discussed in connection with the monotonic measurement constraint of binominal *each*, the monotonicity constraint in (53-b) can be satisfied while the same constraint in (54-b) cannot. This is because extensive measurement yields degree names that track the mereological structure of the distributivity key, i.e., the waters in this case. However, intensive measurement fails to yield degree names that have the same effect.

Counting quantifiers are taken to contribute both individual d-refs and degree d-refs. If all *ge* needs is one d-ref (whichever d-ref) to satisfy its monotonicity constraint, then counting quantifiers are predicted to be acceptable, as long as the individual d-ref they contribute co-varies with distributivity. Since degree names are derived from measuring individuals, if a counting quantifier does not contribute individuals that co-vary with distributivity, then the degree names derived from measuring these individuals will also remain constant relative to distributivity, in violation of the definition of the monotonicity constraint in (45).

5.3.2 Failure of licensing by bare noun phrases

Recall that *ge* is unacceptable when the distributed share only contains a transitive verb and a bare noun phrase.

(55) *Zhe-xie haizi ge kan-le shu.
this-CL.PL child GE see-ASP shu
'*The children read (the) books each.'

Bare noun phrases may receive a definite interpretation or an existential, indefinite interpretation in Mandarin (Cheng and Sybesma 1999, Yang 2001, Trinh 2011, Dayal 2011a, Jiang 2012). As discussed earlier, both interpretations fail to license *ge*, for similar reasons. I first discuss why the existential interpretation fails to satisfy the monotonicity constraint, and then move on to discuss why the definition interpretation also fails to do so.

We have seen, in Chapter 4, that existentially interpreted bare noun phrases can satisfy the independence constraint of Cantonese *saai*.[9] Based on the interaction of bare noun phrases and *saai*, I have argued that existentially interpreted bare noun phrases are kind terms and they only introduce a kind-level d-ref (see also Cheng and Sybesma (1999) and Jiang (2012), who argue that bare noun phrases in Cantonese are kinds). The existential interpretation comes from the interpretation procedure, which employs a mechanism akin to Derived Kind Predication (Chierchia 1998, see also Carlson 1977a).[10]

[9] According to Cheng and Sybesma 1999, bare noun phrases in Cantonese cannot receive a definite interpretation (see also Jiang 2012). So, it is the existential interpretation of bare noun phrases in Cantonese that satisfy the independence constraint.

[10] Farkas and de Swart (2003) also uses a similar interpretive mechanism to derive the existential force of a semantically incorporated bare singular (i.e., bare noun phrase that lacks plural morphology), which they analyze as uninstantiated arguments rather than kinds.

I will assume that an existentially interpreted bare noun phrase in Mandarin is also treated as a kind term (see also Yang 2001, Trinh 2011, Jiang 2012). Since a kind term is similar to a proper name and does not co-vary with distributive quantification, a bare noun phrase like *shu* 'books' receives a similar translated as a proper name:

(56) $\textbf{books}^{y^k} := \lambda P.\exists y^k \wedge y^k = \text{bk-kind} \wedge P\,y$

(57) $\textbf{John}^{y} := \lambda P.\exists y \wedge y = \text{j} \wedge P\,y$

The kind-level individual bk-kind is modeled, following Chierchia (1998), as a function from a world to a plurality consisting of all the instantiations of the book kind in that world:[11]

(58) $\text{bk-kind} := \lambda s.\iota x.\text{bks}\;x$

Suppose the possible values for y^k are drawn from the domain of kinds D_k.[12] Without any modification, a plain bare noun picks out the biggest plural individual satisfying the NP property (i.e., bk-kind) and this value is stored in y^k, as shown in (56).

The kind-based analysis predicts that bare noun phrases do not license *ge*, because the kind-level d-refs introduced by bare noun phrases are just like proper names and cannot induce co-variation. Concretely, translating (55) into DPlLM gives rise to (59). Interpreting (59) against a set of input info-states yields a set of output info-states. A member in the output is given in Figure 5.3 as an illustration.[13] In this info-state, there is no co-variation between the children and the kind of object they read, as they all read the same kind of object, namely, books. If there is any additional info-state in the output, they all share the same feature as this info-state: y^k stores a unique kind that fails

[11]A intensional compositional semantics is needed to compose a kind term and a predicate.

[12]Members in D_k are derived by applying a nominalization operator $^\cap$ to a predicate (Chierchia 1984, 1998). According to these studies, $^\cap P$ returns a function from an index to the maximal individual having the property P in that index.

[13]Since maximization dose not guarantee uniqueness, the output set here will have more than one info-state if distributive quantification is assumed to allow non-atomic distributivity (also known as cover-based distributivity). For example, another info-state in the output can have the following form:

H	x	y^k
h_1	c1⊕c2	bk-kind
h_3	c3	bk-kind

If non-atomic distributivity is ruled out, then H will be the only output. A requirement I have not discussed in this study is the fact that *ge* strongly favors atomic distributivity, in contrast to *dou*, which is compatible with both atomic and non-atomic distributivity (Lin 1998a.)

to co-vary with distributivity.

(59) $\text{max}^x(\text{child } x) \wedge \delta_x(\exists y^k \wedge y^k = \text{book-kind} \wedge \text{read } y^k\ x) \wedge \text{dm}_{x,y^k}$

H	x	y^k
h_1	c1	bk-kind
h_2	c2	bk-kind
h_3	c3	bk-kind

...

Figure 5.3: An sample info-state after interpreting (59)

I follow Trinh (2011) and assume that when a bare noun phrase receives a definite interpretation, there is an extensional operator **EXT** that saturates the world argument of a relevant kind term and as a result yields a maximal individual (not an individual concept) that is a member of the kind. Accordingly, a bare noun phrase with a definite interpretation is treated in the same way as a definite noun phrase, which denotes a dynamic generalized quantifier:

(60) $\textbf{EXT books}^y := \lambda P.\text{max}^y(\text{book } y) \wedge P\ y$

This quantifier introduces a maximal individual that has the book property. This individual, even if interpreted inside the scope of distributivity, does not co-vary with distributivity, as shown in (61). For this reason, the monotonicity constraint of *ge* cannot be satisfied by a bare noun phrase receiving a definite interpretation.

(61) $\text{max}^x(\text{child } x) \wedge \delta_x(\text{max}^y \wedge (\text{book } y) \wedge \text{read } y\ x) \wedge \text{dm}_{x,y}$

Given the contrasting requirements of the monotonicity constraint of *ge* and the independence constraint of Cantonese *saai*, the fact that bare noun phrases pattern differently with respect to these two constraints is a welcome result.

5.3.3 Licensing by bound pronouns

I have suggested that allowing *ge*'s monotonicity constraint to be satisfied by either a dependent degree variable or a dependent individual variable suffices to account for all of *ge*'s licensing conditions that do not involve a measurement component. In this section, I show why bound pronouns can license *ge*, as exemplified in (62) (repeated from (21)):

(62) Zhe-xie haizi^x ge kan-le ziji$_x$/*Zilu dailai-de shu$^{y/y^k}$.
these-CL.PL child GE read-ASP self/Zilu bring-DE book
'These children each read the book(s) they/??Zilu brought.'

To interpret this sentence, we need to decide how to interpret complex bare noun phrases like *ziji$_x$-dailai-de shu* 'books x bought' and *Zilu dailai-de shu* 'books Zilu brought'. It turns out that we cannot make a uniform decision for all bare noun phrases with a relative clause modifier. Some of them should be treated as definite expressions while others should be treated as kinds. Bare noun phrases like *Zilu dailai-de shu* 'books Zilu brought' should not be treated as kinds, as they are incompatible with kind-level predication, as shown in (63), unlike their unmodified counterparts, as shown in (64).

(63) #Zilu dailai-de shu juezong le.
Zilu bring-MOD book extinct SFP
'Books that Zilu brought are extinct.'

(64) Shu juezong le.
book extinct SFP
'Books are extinct.'

A context in which (64) is acceptable is when paper books become completely replaced by electronic texts, or when there is a new way to transmit knowledge that does not rely on the use of books of any form. I cannot think of a context in which (63) can be used felicitously. A kind is an individual concept, which still exists even when there is no instantiation of the kind in a particular world (Chierchia 1998). This is why (64) is acceptable. However, *books Zilu brought* does not have enough trans-world instantiations to form an individual concept. In other words, there is no concept independent of the actual world instantiation of the entities. This is why (63) is incompatible with a kind-level predicate.

Based on the contrast in (63) and (64), I suggest modeling *ziji$_x$-dailai-de shu* 'books x bought' and *Zilu dailai-de shu* 'books Zilu brought' as definite expressions:

(65) ι **books-x-bring** := $\lambda P.\mathsf{max}^y(\mathsf{book}\ y \wedge \mathsf{bring}\ y\ x) \wedge P\ y$

(66) ι **books-Zilu-bring** := $\lambda P.\mathsf{max}^y(\mathsf{book}\ y \wedge \mathsf{bring}\ y\ \mathsf{z}) \wedge P\ y$

Note that when the bare noun phrase contains a pronoun (modeled as a variable), as in the case of (66), the maximal individual picked out by the definite expression and introduced into an info-state depends on the value of the pronoun. So, when the pronoun co-varies with the distributivity key, the maximal individual introduced into an info-state also may (but need not) co-vary with the distributivity key, as shown in (67-b). For this reason, the monotonicity constraint of *ge* can be satisfied in (67-a) (a nemonic for the corresponding Mandarin sentence in (62)).

(67) a. The children *ge* read books <u>they</u> brought.

b. $\mathsf{max}^x(\mathsf{child}\ x) \wedge \delta_x(\mathsf{max}^y(\mathsf{book}\ y \wedge \mathsf{bring}\ y\ x) \wedge \mathsf{read}\ y\ x) \wedge \mathsf{dm}_{x,y}$

However, when the pronoun is replaced by a proper name, co-variation is no longer available. This is because the individual picked out by a proper name cannot co-vary with the distributivity key. As a result, the maximal books that this individual brought also cannot co-vary with the distributivity key, as shown in (68-b). This explains why (68-a) is unacceptable.

(68) a. ??The children *ge* read books <u>Zilu</u> brought.

b. $\mathsf{max}^x(\mathsf{child}\ x) \wedge \delta_x(\mathsf{max}^y(\mathsf{book}\ y \wedge \mathsf{bring}\ y\ \mathsf{z}) \wedge \mathsf{read}\ y\ x) \wedge \mathsf{dm}_{x,y}$

Now, we can take up a different type of bare noun phrases modified by a relative clause, i.e., those that may assume a kind interpretation. Discussing them should allow us to see that the contribution of a bound pronoun is neutral to whether bare noun phrases are analyzed as kinds or definite expressions in Mandarin. To begin with, observe that a bare noun phrase like *Zilu xihuan-de dongwu* 'animals that Zilu likes' is compatible with a kind-level predicate:

(69) Zilu xihuan-de dongwu juezhong le.
Zilu like-MOD animal extinct SFP
'Animals that Zilu likes are extinct.'

Then, observe that this kind of bare noun phrase can license *ge* when the relative clause contains a pronoun but not when the pronoun is replaced by a proper name:

(70) Zhe-xie haizi^x ge kanjian-le ziji$_x$/*Zilu xihuan-de dongwu$^{y^k}$.
these-CL.PL child GE see-ASP self/Zilu like-DE animal
'These children each saw the animal(s) they/??Zilu like.'

If we take the complex noun phrases as a kind term, then the whole noun phrase can be translated as a generalized quantifier (following the analysis of proper names in dynamic semantics). The meaning of *Zilu dailai-de shu* 'books brought by Zilu' as well as the meaning of a plain, unmodified bare noun phrase are given below for comparison:

(71) **animals-x-like** := $\lambda P.\exists y^k \wedge y^k =$ animals-x-likes) $\wedge P\ y^k$

(72) **animals Zilu like** := $\lambda P.\exists y^k \wedge y^k =$ animals-z-like $\wedge P\ y^k$

The difference between (71) on the one hand and (72) on the other hand lies in whether y^k is allowed to store values that may co-vary with another variable. To begin with, recall that a kind is formed by applying the nominalization operator $^\cap$ to a predicate (Chierchia 1984, 1998). If the predicate is complex, as in the case of *ziji dailai-de shu* 'books he/she brought' and *Zilu dailai-de shu* 'books Zilu brought', I assume that $^\cap$ applies after the modification. From the definitions below, we can see what difference having a variable in the modifier makes:

(73) animals-x-likes = $^\cap\lambda y.$*animal $y\ \wedge$ *like $y\ x$

(74) animals-zilu-likes = $^\cap\lambda y.$*animal $y\ \wedge$ *like y zilu

In (73), the predicate being nominalized may have different members depending on the value of x. As a result, the kind derived from this predicate may also vary depending on the value of x. However, in (74), the predicate subject to nominalization is a fixed set, i.e., the set of things that are books and brought by Zilu. Accordingly, the kind derived from this predicate is also fixed. For this reason, in (71) y^k can (but need not) store different kinds of books depending on the values assigned to x. However, in (72) y^k can only store a single kind of books, namely, books that Zilu brought. In this sense, a bare noun with a modifier that does not contain a variable, as in the case of (72), is similar to a plain, unmodified bare noun phrase, since both of them do not induce co-variation.

The contrast can be more clearly illustrated when (71) and (72) are used in a sentence like (70). The corresponding translations are given below (the English sentences are included as mnemonics for the corresponding Mandarin sentences):

(75) a. The children *ge* read books <u>they</u> brought.

b. $\mathsf{max}^x(\mathsf{child}\ x) \wedge \delta_x(\exists y^k \wedge y^k = \mathsf{animals}\text{-}x\text{-}\mathsf{like} \wedge \mathsf{see}\ y^k\ x) \wedge \mathsf{dm}_{x,y^k}$

(76) a. ??The children *ge* read books <u>Zilu</u> brought.

b. $\mathsf{max}^x(\mathsf{child}\ x) \wedge \delta_x(\exists y^k \wedge y^k = \mathsf{animals\text{-}z\text{-}likes} \wedge \mathsf{see}\ y^k\ x) \wedge \mathsf{dm}_{x,y^k}$

In (75-b), y^k may vary depending on the value assigned to x. For this reason, the monotonicity constraint can be satisfied. By contrast, in (76-b), y^k cannot vary relative to x and hence the monotonicity constraint cannot be satisfied.

Before leaving this section, it is useful to point out that if a pronoun is not bound by the distributivity key, it is a lot trickier to determine whether a noun phrase containing the pronoun co-varies with the distributivity key. Typically, a free pronoun is not allowed, as reported earlier in Section (30). The relevant example is repeated below.

(77) Zhe-xie haizi^x ge kan-le tamen$_{x/*y}$-de shu.
these-CL.PL child GE read-ASP they-DE book
'These children each read their$_{x/*y}$ book.' Bound/*Free

However, if there is enough contextual support for a dependency between the distributivity key and the plural individual referred the plural pronoun refers to, a similar sentence can be improved:

(78) a. *Context:* Each child x asked a friend y to bring a book for them.

b. Zhe-xie haizi^x ge kan-le tamen$_{x/?y}$ dailai-de shu.
these-CL.PL child GE read-ASP they bring-DE book
'These children each read books they$_{x/?y}$ brought.'

The improved judgment is not entirely unexpected. Although the pronoun is not directly bound by the distributivity key, it does co-vary with the distributivity. In particular, the pronoun refers to a different friend for each child. Since the friends stand in a dependency with the children, the books the friends brought also co-vary with the children, too.

Sentence-internal adjectives

To understand why sentence-internal readings of *butong* 'different' licenses *ge*, it is necessary to delve deeper into the semantics of *butong*. I am not aware of any semantic studies on Mandarin *butong*, but it is possible to borrow insights from research on English *different*.

Many studies distinguish between two variants of *different*. One variant is 'singular *different*' (as in *a different poem*) and the other is 'plural *different*' (as in *different poems*) (Beck 2000, Brasoveanu 2011, Bumford 2015). Both singular and plural *different* have an external interpretation as well as an internal interpretation (see Section 5.2.2 for distinguishing between the two). Since only the internal interpretation licenses *ge*, it is our main focus here. As a clarification, all the sentences and analyses given in this section are to be understood as targeting the internal interpretation of *butong* and *different*. The motivation for distinguishing between two variants of *different* comes from their distinct distributions (I refer the reader to a detailed discussion of their distributions in Brasoveanu (2011)). In particular, singular *different* has a more restricted distribution than plural *different*, as reported in Brasoveanu (2011). A small paradigm showing the difference is given below:

(79) Every boy recited a different poem / different poems. Brasoveanu 2011: (7), (24)

(80) The boys recited different poems / #a different poem. Brasoveanu 2011: (19), (26)

Noun phrases in Mandarin do not have number morphology. However, Mandarin also seems to distinguishe between two types of *butong*. Let us call them 'singular' and 'numberless' *butong*. First, observe that the two types of *butong* also have distinct distributions. Numberless *butong* is compatible with plural predication (81) as well as distributivity, brought about by the distributivity marker *dou*, as in (82), or a combination of *meige*-NP 'every-NP' and *dou*, as shown in (83). However, singular *butong* is only fully acceptable when there is a singular distributive quantifier, as in the case of (83). It is degraded in other environments.

(81) Zilu he Ziyou kan-le (??yi-ben) butong-de shu.
Zilu and Ziyou read-ASP one-CL different-MOD book
'Zilu and Ziyou read different books.'

(82) Zhe-xie haizi dou kan-le (?yi-ben) butong-de shu.
this-CL.PL DOU read-ASP one-CL different-MOD book
'These children all read different books.'

(83) Meige haizi dou kan-le (yi-ben) butong-de shu.
every child DOU read-ASP one-CL different-MOD book
'Every child read a different book / different books.'

As far as *ge* is concerned, it is compatible with both singular and numberless *butong* but in distinct environments. When the subject is plural, both singular and numberless *butong* can be used, as shown in (84). When the subject is singular and distributive, as in the case of *mei-ge NP*, 'every child', singular *butong* is strongly preferred, as shown in (85).

(84) Zhe-xie haizi ge kan-le (yi-ben) butong-de shu.
this-CL.PL child GE read-ASP one-CL different-MOD book
'These children each read a different book / different books.'

(85) Meige haizi ge kan-le ??(yi-ben) butong-de shu.
every child GE read-ASP one-CL different-MOD book
'Every child read a different book / different books.'

The distribution of *butong* is summarized in Table 5.4. I do not try to offer an account for the distribution of *butong* here, other than noting that singular *butong* prefers environments in which atomic distributivity (distributivity down to atoms) is required.

Meige-NP + *ge*	SG
Plural + *ge*	SG, Num-less
Meige-NP + *dou*	SG, Num-less
Plural + *dou*	?SG, Num-less
Plural	Num-less

Figure 5.4: The distribution of singular and numberless *butong* in Mandarin

Extant studies on English *different* have largely assumed that singular and plural *different* have distinct, albeit related semantics. In particular, singular *different* is argued, in Beck (2000), Brasoveanu (2011), and Bumford (2015), to be a 'quantifier-internal' phenomenon. The licensing of singular *different* crucially relies on the analysis of singular distributive quantifiers. In Beck (2000) and Brasoveanu (2011), singular distributive quantifiers involve universal quantification over pairs (of individuals in Beck (2000) and of info-states in Brasoveanu (2011)). In Bumford (2015), singular distributive quantifiers are analyzed as involving incremental quantification. By contrast, plural *different* is modeled as involving cumulative predication and reference to dependencies arising from it (pragmatically with use of covers in Beck (2000) and semantically with use of PCDRT in Brasoveanu (2011)).[14]

Since universal quantification and distributivity in Mandarin and English differ considerably, it is not surprising that singular and numberless *butong* do not line up neatly with singular and plural *different*. However, given the distinct distributions of singular and numberless *butong*, it is imaginable that an adequate theory of them needs to distinguish between different modes of distributive quantification. However, for the purpose of understanding the licensing of *ge*, I will make the simplifying assumption that singular and numberless *butong* only differ with respect to the presence of a counting component. They are licensed by the same mode of distributive quantification. Given this assumption, I will not develop a new theory of universal quantification to accommodate the differences between singular and numberless *butong*.

My analysis of *butong* follows largely the spirit of Kuhn (2017)'s analysis of English *same*. More specifically, I assume, following Kuhn (2017), that sentence-internal adjectives are like dependent indefinites, in the sense that they can receive delayed evaluation. To model the delayed evaluation, I assume that noun phrases containing a sentence-internal adjective receives a higher-order meaning, along the lines of Cresti (1995), de Swart (2000), Charlow (to appear), and the analysis developed for binominal *each* in Chapter 3.

Concretely, let me start by translating a noun phrase with a singular *butong*, such as *yi-ben butong-de shu* 'a different book'. The numberless *butong* involves more complexities and will be dealt with after we walk through singular *butong*. The translation is offered in (86), and the meaning

[14] In addition, Barker (2007) suggests a parasitic scope analysis for plural *different*, which is an extension of the analysis developed for *same*. Kuhn (2017) makes suggestive notes about the semantics of *different* based on the semantics of *same*, but does not actually offer a concrete formulation.

of of the non-identity condition $\mathsf{diff}_{x,y}$ is given in (87).

(86) $\textbf{a different}_{x,y}\ \textbf{book}^{y} := \lambda c.c(\lambda P.\exists y \wedge \mathsf{book}\ y \wedge |y| = 1 \wedge P\ y) \wedge \mathsf{diff}_{x,y}$

(87) $G[\![\mathsf{diff}_{x,y}]\!]H = \mathbb{T}$ iff for all $a, b \in G(x) : G|_{x=a}(y) \neq G|_{x=b}(y)$

Note first that in (86) x is the variable being distributively quantified. It receives a value when the rest of the sentence saturates the c argument. y is the variable subject to the non-identity relation. These variables allow *butong* to gain access to the dependencies arising from distributive quantification. More importantly, the test $\mathsf{diff}_{x,y}$ is introduced outside the scope of distributivity, which is introduced inside c and scopes over no more than the existential quantifier contributed by *a different book*. The truth condition of $\mathsf{diff}_{x,y}$ makes sure that all the true transitions return an info-state in which for each pair of values in x, the corresponding y stores distinct values. For concreteness, (88-b) represents the translation of a sentence involving singular *butong*:

(88) a. The children *ge* read a different book.

b. $\mathsf{max}^{x}(\mathsf{child}\ x) \wedge \delta_x(\exists y \wedge |y| = 1 \wedge \mathsf{book}\ y \wedge \mathsf{read}\ y\ x) \wedge \mathsf{diff}_{x,y} \wedge \mathsf{dm}_{x,y}$

As can be seen above, the 'quantifier-internal' property of singular *different* is gone. Access to the information outside the scope distributive quantification is done strictly outside the scope of distributive quantification, via quantificational subordination.

We now turn to the meaning of a noun phrase with numberless *butong*, such as *butong-de shu* 'different books'. As warned earlier, treating numberless *butong* is not as straightforward as treating singular *butong*. To appreciate the challenge, let us first consider two different ways to translate *butong-de shu* 'different books', (89) and (90). They are not equally good candidates.

(89) $\textbf{different}_{x,y}\ \textbf{books}^{y} := \lambda c.c(\lambda P.\exists y \wedge \mathsf{book}\ y \wedge P\ y) \wedge \mathsf{diff}_{x,y}$

(90) $\textbf{different}_{x,y}\ \textbf{books}^{y^k} := \lambda c.c(\lambda P.\exists y^k \wedge y^k = \mathsf{bk\text{-}kind} \wedge P\ y^k) \wedge \mathsf{diff}_{x,y^k}$

In particular, (89) offers a good interpretation while (90) is destined to yield a contradiction. It is easy to see the difference. With (89) y may store individuals that co-vary with x, the variable that will store the values provided by the distributivity key once c is saturated. However, with (90) y^k is specified to be bk-kind and may never co-vary with x, according to our analysis in Section 5.3.2.

What the contrast suggests to us, then, is that when *butong* acts as a modifier of a bare noun phrase *shu* 'books', the whole bare noun phrase cannot resolve to the maximal plural individual having the property of books. Can we then, still maintain a consistent story that bare noun phrases receive a kind interpretation, which is not useful for licensing *ge*? I argue that we can.

Recall that the reason a pronoun can help give rise to variation lies in the fact that the pronoun is inside the scope of the nominalization operator, as repeated in (91). In other words, its contribution makes a difference for kind formation.

(91) $\textsf{bk-kind-}x\textsf{-brought} = {}^{\cap}\lambda y.{}^{*}\text{book } y \wedge {}^{*}\text{brought } y\ x$

Now, consider our definition of (90). We end up with bk-kind as the meaning of *butong-de syu* 'different books' because we have *not* considered the contribution of *butong* in kind formation. In other words, we have not given it the chance we have given to a bound pronoun. The reason why we excluded it so easily in the first place is because of the decision to avoid getting into a more complex semantics for distributive quantification and use delayed evaluation as a means to model quantifier-internal phenomena. Setting aside for the time being how to change our semantics of distributive quantification, let us consider what would be the result had we included the contribution of *butong* at the level of kind formation.

Let's call the kind corresponding to $butong_{x,y}$*-de shu* 'different books' '$\textsf{diff}_x\textsf{-bk-kind}$', where x is the variable subject to distributive quantification (there is no need to use the variable y). Depending on the value associated with x, $\textsf{diff}_x\textsf{-bk-kind}$ returns a different kind of books. For example, if x is assigned a value a, then $\textsf{diff}_x\textsf{-bk-kind}$ is the kind of books that that no one else in $G(x)$ other than a read.

Clearly, to model this type of kind formation, we need access not just to the value assigned to x by a single assignment, but all the other values assigned to x by other assignments, right at the site of kind formation. To provide a semantics to accommodate this is outside the scope of this dissertation. However, the discussion above makes it clear that bound pronouns and sentence-internal adjectives indeed pattern the same in licensing *ge*—they both provide an opportunity for co-variation, something not available for a plain bare noun phrase.

5.3.4 Respective distributivity

In this section, I offer a formulation of respective distributivity in DPlLM based on the analysis of Gawron and Kehler (2004). There are other studies that offer analyses to the respective interpretation, such as Kubota and Robert (2016). However, since modeling the respective interpretation is not the main goal of this chapter, I pick Gawron and Kehler (2004) as the starting point because it shares ontological assumptions with the present study with respect to the presence of referential pluralities, as well as functions that turn referential pluralities to pluralities at the evaluation level.

The respective distributivity operator, as proposed in Gawron and Kehler (2004), has the following form (to avoid notational confusions, I have swapped their g (group) for x, which can a be a variable for plural individuals in the present work):

(92) $$\textsc{resp}_f := \lambda P \lambda x. \bigsqcup_{1 \leq i \leq |f|} [f(P)(i)](f(x)(i))$$ (Gawron and Kehler 2004:(14))

According to Gawron and Kehler (2004), this operator, with a pragmatically available sequencing function f, takes a property sum P, a plural individual x, and returns a proposition sum. The proposition sum is the collection of all propositions that is obtained by applying one property from the property sum to one individual in the plural individual. The application is guided by a sequencing function f. f breaks down a plurality (a plural individual or a property sum) into sub-pluralities (typically atoms) and labels each sub-plurality with a bigger number starting from 1. The set of numbers used for labeling the sub-pluralities is the cardinality of f, i.e., $|f|$. The numbers serve as a guide for f to find a particular sub-plurality. The functional application of one plurality to another plurality is guided by the numerical labels on the two pluralities, such that $f(P)(1)$ is applied to $f(x)(1)$ and $f(P)(2)$ is applied to $f(x)(2)$, so on and so forth. f is additionally required to satisfy the following requirements, according to Gawron and Kehler (2004) (pp.173–174).

(93) Requirements on a sequencing function

a. Same cardinality: all pluralities that serve as arguments to $\textsc{Resp}_f$ must have the same cardinality.

b. Proper subgroups: for each x and i, $f(x)(i)$ picks a proper subpart of x.

c. Exhaustivity: summing up all the sub-pluralities generated by f on x returns x, i.e.,

$$\left(\bigsqcup_{1\leq i\leq |f|} f(x)(i)\right) = x$$

These requirements are intended to explain the restricted distribution of respective distributivity. The same cardinality requirement predicts that a typical cumulative interpretation that lacks information about one-to-one correspondence between the sub-pluralities in two pluralities do not give rise to a respective interpretation:[15]

(94) Five hundred companies used six hundred computers, (*respectively).

The requirement on proper subgroups makes sure that a plurality and a singleton do not license the respective interpretation:

(95) John and Mary saw Peter, (*respectively).

Lastly, the exhaustivity requirement rules out cases in which f does not pick out all the parts in a plurality. For example, if there is an f that only picks out, for (96), *John* from the plurality *John and Mary*, and the property *jogged* from the property sum *jogged and swam*, then this f is not usable for respective distributivity because it fails exhaustivity.

(96) John and Mary jogged and swam, respectively.

Given the definition in (92) and the requirements in (93), the following sentence is evaluated as in Table 5.1.[16]

---

[15]Determining the cardinality of a plurality is not as straightforward as just finding out all the atoms in the plurality. Gawron and Kehler (2004) discuss pluralities involving duplicate parts, such as the coordinated VP in (i-a) below. If pluralities are treated as sets or sums, then the coordinated VPs in (i-a) and (i-b) have the same cardinality, namely, 2. However, the fact that (i-a) is acceptable while (i-b) is not suggests that some way is needed to model pluralities with duplicate parts. In this study, I assume, along the lines of Gawron and Kehler (2004), that duplicates can be represented, but remain open as to how to model them. One possibility is discussed in Kubota and Robert (2016), who model pluralities using multisets, i.e., sets that allow for duplicate occurrences of identical elements.

(i) a. Sue, Karen, and Bob jog, drive, and jog respectively.
b. #Sue, Karen, and Bob jog and drive respectively.

[16]Given that a sequencing function is pragmatically determined, (97) in principle allows a different f that pairs Zilu with dancing and Ziyou with singing. I believe this is true, as (i) is not a contradiction.

(97) Zilu he Ziyou fenbie chang-le ge he tiao-le wu.
Zilu and Ziyou respectively sing-ASP song and dance-ASP dance
'Zilu and Ziyou sang and danced, respectively'

	Individual sum	Property sum
	$f(\mathsf{zl} \oplus \mathsf{zy})(i)$	$f(\mathsf{sing} \oplus \mathsf{dance})(i)$
1	zl	sing
2	zy	dance

Table 5.1: Respective distributivity as in Gawron & Kehler (2004)

There are close connections between Gawron and Kehler (2004)'s analysis of respective distributivity and distributivity in a plural logic (DPlL/DPlLM/PCDRT). The sequencing function f plays a similar role as a set of assignments. In Gawron and Kehler (2004), f splits up a plurality, whereas in a plural logic, a set of assignments splits up a plurality. In Gawron and Kehler (2004), f establishes a correspondence relation between parts in two pluralities, whereas the same job is tasked to a set of assignments in a plural logic. Lastly, the correspondence relation establish by f allows an evaluation to proceed pair by pair, giving rise to distributivity. In a plural logic, since a set of assignments are used to store a correspondence relation, distributivity is achieved by splitting up the set of assignments. Lastly, both frameworks returns a plurality as the result of respective distributivity. In Gawron and Kehler (2004), the result is a proposition sum (i.e., a list of propositions), whereas in a plural logic, the result is a set assignments, which store information that can verify a proposition sum. Based on these similarities, Gawron and Kehler (2004)'s insights can be straightforwardly translated into DPlLM.

To ease into the discussion, first consider a sentence with two coordinate noun phrases under a simple cumulative interpretation.

(98) Zilu he Ziyou kanjian-le Zixia he Zisi.
Zilu and Ziyou see-ASP Zixia and Zisi

(i) Zilu he Ziyou fenbie chang-le ge he tiao-le wu. Zilu tiao-de wu, Ziyou chang-de ge.
Zilu and Ziyou respectively sing-ASP song and dance-ASP dance but I not-know who do-ASP what
'Zilu and Ziyou sang and danced, respectively. Zilu danced and Ziyou sang.'

'Zilu and Ziyou saw Zixia and Zisi, respectively'

As suggested in Section 2.3 of Chapter 2, this sentence receives the following interpretation:

(99) $\mathsf{max}^x(x = \mathsf{zl} \oplus \mathsf{zy}) \wedge \mathsf{max}^y(y = \mathsf{zx} \oplus \mathsf{zs}) \wedge \mathsf{saw}\ y\ x$

The outcome is a collective/cumulative interpretation, which does not encode any dependency between x and y. A respective interpretation differs minimally from a collective/cumulative interpretation with the addition of a respective distributivity operator. The respective distributivity operator has flexible arity. In (113), I assume that it is adjoined to the transitive verb and takes the verb as one of its arguments. It takes two other arguments, namely, a pair of d-refs storing two pluralities. These two d-refs will be used to construct a pair of new d-refs subject to distributive evaluation. When a coordinated VP is involved, the respective distributivity operator then only has two arguments, i.e., a pair of d-refs. One d-ref stores a plural individual and the other stores a property sum, following Gawron and Kehler (2004)'s study.

(100) $\mathsf{max}^x(x = \mathsf{zl} \oplus \mathsf{zy}) \wedge \mathsf{max}^y(y = \mathsf{zx} \oplus \mathsf{zs}) \wedge \mathsf{Resp}_{x,y}^{x',y'}(\mathsf{saw}\ y\ x)$

The definition of $\mathsf{Resp}_{x,y}^{x',y'}(\phi)$ is given in (101). It can be divided into two parts: a pair-wise variable introduction, notated as $\mathsf{Resp}_{x,y}^{x',y'}$ and defined in (101-a), and a distributive evaluation of ϕ, notated as $\delta_x(\phi)$ and defined in (101-b).[17] I spell these two parts in turn below.

The pair-wise variable introduction introduces two new d-refs x' and y' based on two extant variables x and y and a sequencing function f as defined in Gawron and Kehler (2004). x and y are independently introduced with use of two coordinated noun phrases. They can, but need not, stand in any dependence relation. The variables x' and y', however, are *not* introduced in by the default variable introduction, i.e., $\exists x' \wedge \exists y'$. In particular, (omitting the subscripted d-refs for

[17] In sentences with more than two coordinated noun phrases that exhibit respective distributivity, such as (i), pair-wise variable introduction will need to be generalized to tuple-wise variable introduction.

(i) John and Mary wanted to give a book and a pen to Sue and Jane, respectively.

the moment) with $\mathsf{Resp}^{x',y'}$ each assignment g in the input G simultaneously updates, in a pair-wise manner, the x' slot and the y' slot using the pairing information provided by the sequencing function. As a result, x' and y' do stand in a dependence relation, as long as x and y store proper pluralities. By assuming that this special variable introduction rule is available only when there is a salient sequencing function inferable from the context has the effect of allowing a respective interpretation only when all the requirements on the sequencing function is met.

The distributive evaluation of ϕ is facilitated by the good old distributivity operator $\delta_{x'}$ we have been using throughout the dissertation.[18] It splits up the evaluation along the x'-dimension and checks that in each sub-info-state storing one x'-value, evaluation of ϕ leads to at least one output info-state.

(101) $G[\![\mathsf{Resp}^{x',y'}_{x,y}(\phi)]\!]H := \mathbb{T}$ iff (a) and (b) below

a. There is H' such that $H' = \{g^{x' \to f(x)(i), y' \to f(y)(i)} | \, g \in G \;\&\; f(x)(i) <_A G(x) \;\&\; f(y)(i) <_A G(y)\}$, where $<_A$ is the atomic part-of relation;

b. For all $a \in H'(x).H'|_{x=a}[\![\phi]\!]H|_{x=a}$

A graphical illustration of the interpretation of (100) is given in Figure 5.5.

As one can see, after evaluating (100), any info-state in the output will store two variables that exhibit dependence. It is for this reason that *ge* is licensed by respective distributivity. As shown more concretely in (102), *ge* can use the variables x' and y' introduced by pair-wise variable introduction to satisfy the monotonicity constraint.

(102) $\mathsf{max}^x(x = \mathsf{zl} \oplus \mathsf{zy}) \wedge \mathsf{max}^y(y = \mathsf{zx} \oplus \mathsf{zs}) \wedge \mathsf{Resp}^{x',y'}_{x,y}(\mathsf{saw}\ y\ x) \wedge \mathsf{dm}_{x',y'}$

In short, evaluating respective distributivity gives rise to info-states that encode variable dependence, which *ge* can use to satisfy the monotonicity constraint.

[18] x' is the variable storing the atomic sub-pluralities extracted from the first mentioned coordination. Since x' and y' are introduced in a pair-wise manner and their value stand in one-to-one correspondence, swapping x' with y' does not generally cause any problem, provided here is a way to model duplication.

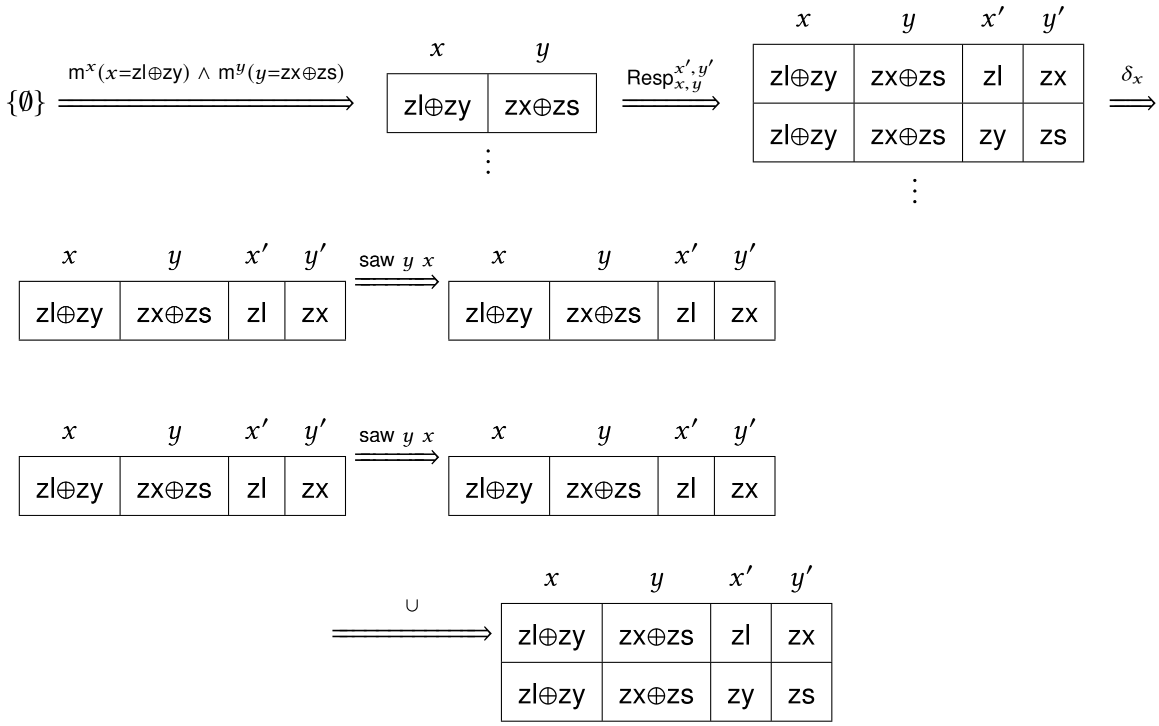


Figure 5.5: Respective distributivity

5.4 Generalized monotonicity with events

In this section, I show that a further generalization of the monotonicity constraint to include events as dependent variables captures the ‘event differentiation’ condition of determiner *each* as discussed in Vendler (1962), Tunstall (1998) and Brasoveanu and Dotlačil (2015).

To begin with, Vendler (1962) notes that there is a difference between (103) and (104):

(103) Take every apple.

(104) Take each apple.

He notes that with (103) the speaker doesn’t care how the apple is being taken. However, with (104) the speaker intends for the apple to be taken one by one. This contrast has been taken up further in Beghelli and Stowell (1997) and Tunstall (1998) as evidence that *every* and *each* have distinct grammatical properties. Tunstall (1998) posits an event differentiation condition to distinguish between quantification with *every* and quantification with *each*, as cited in (105):[19]

(105) Differentiation Condition (Tunstall 1998:100):

A sentence containing a quantified phrase headed by *each* can only be true of event structures which are totally distributive. Each individual object in the restrictor set of the quantified phrase must be associated with its own subevent, in which the predicate applies to that object, and which can be differentiated in some way from the other subevents.

Although Tunstall's attempt to experimentally verify the differentiation condition was unsuccessful, Brasoveanu and Dotlačil (2015) were able to verify the presence of the event differentiation condition with improved experiments.

The event differentiation condition, at a descriptive level, is very similar to the variation requirement of binominal *each*. The only difference seems to be the locus of the variation: the variation requirement of binominal *each* is imposed on the individual values associated with its host, while the event differentiation condition of determiner *each* is imposed on the event values, presumably introduced by a verbal predicate interacting with distributivity.

To visualize the similarity, we can turn Tunstall (1998)'s event differentiation condition into an event differentiation constraint accompanying distributive quantification, as shown in (106-b) and interpreted in (106-c):

(106) a. John took each apple.

b. $\mathsf{max}^x(\mathsf{apple}\ x) \wedge \delta_x(\exists e \wedge \mathsf{take}\ e \wedge \mathsf{ag}\ e = j \wedge \mathsf{th}\ e = x) \,\overline{\wedge}\, \mathsf{ed}_{x,e}$

c. $G[\![\mathsf{ed}_{x,e}]\!]H := \mathbb{T}$ iff $G = H$ and $\forall a, b \in G(x) : G|_{x=a}(e) \neq G|_{x=b}(e)$

(106-c) says every assignment that assigns a different apple to x should assign a different value to

[19] *Every NPs* in an object position may receive a non-distributive use, as observed in Schein (1993) and Kratzer (2002). For this reason, the contrast in (103) and (104) does not provide sufficient evidence that distributivity with *every* and distributivity with *each* have distinct properties. A more convincing contrast is provided below:

(i) a. Every student left.
b. Each student left.

(i-a) is compatible with all the students leaving at the same time as a group but (i-b) is not. Since *every NPs* do not give rise to a non-distributive use in the subject position, the contrast above provides better support for the claim that distributivity with *every* and *each* are different.

e. In other words, each apple-taking event is differentiated from the other.

However, at a more technical level, the event differentiation condition is almost always trivially satisfied. Suppose there are three apples: apple1, apple2, and apple3, and John took them all at once. Intuitively, the event differentiation condition is violated. However, (106-b) is not violated. This is because for each apple, there is an apple-taking subevent in which John is the agent and the apple is the theme. If we assume the widely held principle, due to Carlson (1998), that events with distinct thematic participants are distinct events, these apple-taking events are all differentiated, by virtue of having different apples as their themes. In short, event differentiation should be trivially satisfied.

To resolve this problem, Tunstall (1998) suggests instead of comparing event values, we compare the thematic dimensions of the relevant events. Correspondingly, this means an enrichment of the event differentiation constraint along the lines in (107). The contribution of the thematic relation θ, as spelled out in (108), is to relate events with their thematic participants. θ here may be resolved to **agnet**, **theme**, **location**, **temporal trace**, or other relevant thematic dimensions.[20]

(107) $G[\![\mathsf{ed}_{x,e}(\theta)]\!]H := \mathbb{T}$ iff $G = H$ and $\forall a, b \in G(x) : \theta(\bigoplus G|_{x=a}(e)) \neq \theta(\bigoplus G|_{x=b}(e))$

(108) $\theta := \lambda e \lambda x . \theta\ e = x$

We are no longer comparing the values in e, but a thematic dimension of the values in e. If two different events happen to yield the same value along a certain thematic dimension θ, then the event differentiation constraint in (106-c) is satisfied but the version in (107), which appeals to thematic transformation, is violated. So, if John took all the apples in one fell swoop, (106-c) is trivially satisfied, but (107) is only trivially satisfied if θ is set to be **th**, i.e., the thematic dimension that happens to be the distributivity key. Although using the thematic version of event differentiation does not completely avoid the possibility that information coming from a distributivity key helps satisfy the differentiation constraint in a trivial manner, it is a significant improvement. To fix the triviality issue with (107), we can posit a variety of reasons to discourage (or ban) the use of the thematic dimension that also identifies the distributivity key. However, it is unclear how to fix (106-c) without some major reconceptualization of events.

[20] I have not spelled out how θ is obtained. I discuss a few possibilities and their implications in the next subsection.

The thematic function θ in the event differentiation constraint in (107) plays the same role as a measure function μ_{dim} in the monotonic measurement constraint. Of course, we can turn the event differentiation constraint in (107) into a full-blown monotonicity constraint, in the form of a monotonic thematic constraint, as shown in (109).[21]

(109) **Monotonic thematic constraint**

$$G[\![\mathsf{dm}_{x,e}(\theta)]\!]H := \mathbb{T} \text{ iff } G = H \text{ and}$$
$$\forall A, A' \in G(x).A \subseteq A' \to \theta(\bigoplus G|_{x\in A}(e)) \leq \theta(\bigoplus G|_{x\in A'}(e))$$

This constraint essentially requires that relative to increasing more values stored in x, the corresponding events stored in e should yield a bigger value along a certain thematic dimension. In other words, there is a (strongly) positive correlation between the size of the distributivity key and the size of a thematic dimension of the events in the scope of distributivity.

Recall that the monotonic measurement constraint of binominal *each* also requires a (weakly) positive correlation, one that is between the size of the distributivity key and the measurement of the host. Given the parallelism in the monotonic measurement constraint and the monotonic thematic constraint, it is not at all surprising why the same form *each* is used to bear these constraints.

A dynamic analysis of determiner and adverbial *each* not only aligns these two uses of *each* with its binominal use, but also makes additional welcome predictions. For example, according to Beghelli and Stowell (1997: 95), determiner *each* cannot be interpreted as having wide scope relative to a negated VP.[22]

(110) ??Each boy didn't leave.

For Beghelli and Stowell (1997), this is because the event variable is bound by the negative quantifier contributed by negation and unavailable to move to a position for the distributed share. In this study, the ill-formedness arises from the failure of dynamic binding—the monotonic thematic

[21]Here, I'm using a strictly increasing version of monotonicity to make it more in line with Tunstall (1998)'s original event differentiation condition. A weaker version may be adopted if it turns out that the some degree of overlapping is tolerated. For example, if *John took each* is compatible with a scenario in which John took all of the apples but different in batches.

[22]They remarked that certain intonations, presumably those that give negation wide scope, could ameliorate the markedness. This is similar to the case of binominal *each* discussed in Section 3.5 of this chapter: When negation scopes above distributivity, the constraint is exempt from dynamic binding failure.

constraint associated with *each* wants to make sure that a thematic dimension of an event inside the scope of distributivity is monotonic relative to the size of the distributivity key. However, negation is static in nature and does not allow dynamic binding information arising from the event variable to escape from its scope. For this reason, the constraint fails because it cannot access the event variable inside the scope of negation.

A question to be addressed more fully is the source of θ, which is an important ingredient for the monotonic thematic constraint. There are a few possibilities. First, the thematic dimension may be accessed *syntactically*, in a similar way as binominal *each* accesses a measure function. To make this work, we need to assume that thematic roles are syntactically represented, as done in Kratzer (1996) for the **agent** role and in Pylkkänen (2008) for applicative roles. Champollion (2017) generalizes this to more traditional thematic roles such as **theme**, **location**, and **goal**. We also need a way to syntactically access the thematic roles. The fact that determiner and adverbial *each* need not be syntactically adjacent to the **theme** role, which is typically realized by the object, adds considerable difficulty to this approach. Alternatively, the relevant thematic dimension may be obtained *pragmatically*. Whenever there is an event, it is presupposed that it comes with some thematic dimensions. The presupposition can then be tapped into for providing a thematic dimension.

Regardless of which option is chosen, a prediction is that events with less accessible thematic dimensions support *each* less well than events with more accessible thematic dimensions. For example, it has been argued that individual-level predicates either do not have a well-defined spatio-temporal dimension (Kratzer 1995), or their spatio-temporal dimensions are bound by a generic operator (hence do not co-vary with distributivity, Chierchia 1995). It is hence predicted that individual-level predicates, such as *knows the secretary* in (111-b), fare worse with *each* than stage-level predicates, such as *saw the secretary* in (111-a).

(111) a. Each student saw the secretary.

b. ?Each student knows the secretary.

By contrast, since *every* does not have an event differentiation constraint, it is predicted to be compatible with both types of predicates:

(112) a. Every student saw the secretary.

b. Every student knows the secretary.

5.5 Previous studies on Mandarin *ge*

5.5.1 Tsai (2009)

There are a good number of studies that take up *ge*'s licensing conditions, including Kung (1993), Lin (1998b, 2005), Tsai (2009), and Lee et al. (2009a). In this section, I briefly review the main claims in these studies and compare them with the present study.

The analysis proposed in Tsai (2009) is closest to the present one. In that study, *ge* is modeled as a special distributivity operator (called a summation operator), which has a semantics based on RESP_f, the distributivity operator that gives rise to respective distributivity in Gawron and Kehler (2004), and a differentiation condition akin to that of English determiner *each* (Tunstall 1998). As described earlier, distributivity with *respectively* is implemented with RESP_f, the respective distributivity operator. f is a sequencing function that assigns a series of numbers to atomic parts in a plural. The differentiation condition, according to Tsai (2009), ensures that a plurality has more than one atomic part. This is meant to model the variation condition of *ge*. However, given that RESP_f comes with a sequencing function and the sequencing function comes with a proper subpart requirement (the function can only extract a proper part of a sum), the differentiation condition is unnecessary for Tsai (2009). For this reason, Tsai's analysis of *ge* needs only RESP_f not the differentiation condition.

As RESP_f, *ge* takes two arguments, a plural NP and a plural VP, modeled as a property sum. As illustration, (113) can be translated as in (114), according to Tsai's analysis. Co-variation (or differentiation, in Tsai's term), is modeled as a presupposition that the VP argument must be a proper sum. If the presupposition is passed, then RESP_f pairs each individual in the plural NP with a property in the property sum, and sums up all these pairings. The result is a proposition sum, which is very similar to the use of a plural info-state for storing information generated from distributive quantification.

(113) %Zilu he Ziyou ge chang-le ge he tiao-le wu.

Zilu and Ziyou GE sing-ASP song and jump-ASP dance

'Zilu and Ziyou sang a song and performed a dance, respectively.'

(114) RESP_f (sing $\sqcup$ dance) (Zilu $\oplus$ Ziyou)

There are a few important differences between Tsai's analysis and the analysis put forward here. The first one concerns the empirical basis on which the two analyses are built. In the version of Mandarin Tsai (2009)'s study is conducted, sentences like (115) are acceptable. However, in the version of Mandarin I work on, sentences of this form are judged quite marginal, if not entirely unacceptable. This is likely due to a dialectal variation.

(115) Xuanshou men ge dida le zhongdian.
contestant PL each reach ASP finish.line
'The contestants each got to the finish line.' (Tsai 2009: 140)

Based on the acceptability of data like (115), Tsai further argues that telicity is an important factor for determining whether or not a VP can be turned into a property sum. However, in the present study, telicity does not play any role.

The second difference lies in the timing for checking co-variation/differentiation. For Tsai, it is checked before distributive quantification happens. For the present analysis, it is checked after it. This makes distinct predictions on what types of expressions can satisfy co-variation/differentiation. For the present analysis, even if there is no conjunction in a VP, as long as there are eligible expressions that may co-vary with distributivity, such as indefinites, bound pronouns, or adjectives exhibiting a sentence-internal interpretation, a plurality can be formed to satisfy co-variation.

This is not the case for Tsai (2009). When a VP does not provide any conjunction, as in the case of (116), it has to be assumed a special type shifter shifts a singular VP into a plural one, as shown in (117).

(116) Zilu he Ziyou ge du-le yi-ben shu
Zilu and Zilu GE read-ASP one-CL book
'Zilu and Ziyou each read a book.'

(117) read(one.book) $\sqcup$ read (one.book)

Although Tsai argues that a VP with a bound pronoun, as the one in (118), is naturally plural, as translated in (119), and hence need not use the type shifter. However, it is unclear to me how the plurality can be derived before the pronoun gets bound by the distributivity key.

(118) Zilu he Ziyou ge du-le ziji-de shu
Zilu and Zilu GE read-ASP self-MOD book
'Zilu and Ziyou each his own book.'

(119) x_1's book $\sqcup$ x_2's book

5.5.2 Lin (1998) and Lee (2009)

Lin (1998b) discusses many similarities in and differences between *dou* and *ge* that go beyond their distinct interactions with expressions in the distributed share. Since the present study only takes up *ge*'s requirements on the distributed share, I only discuss the parts of Lin's analysis that are relevant to these requirements. In terms of empirical scope, Lin only takes up the role of counting quantifiers (called 'extensional entities') in licensing *ge*, so it is unclear whether the analysis can be extended to other licensing expressions such as bound pronouns and sentence-internal adjectives. In terms of analysis, Lin (1998b) take *ge* to be a distributivity operator with three additional functions: (i) requiring the presence of a counting quantifier in the distributed share, (ii) pairing a value provided by the distributivity key with a value provided by a counting quantifier, and (iii) contributing an 'extensionality' requirement.

It seems Lin (1998b) assumes that a counting quantifier can come to contribute a plurality that can be used for pairing. However, he does not discuss explicitly how the plurality is derived. The implicit assumption seems to be that a plurality can be derived by distributive quantification, an idea shared by many studies that model distributivity using a plural logic. By virtue of the fact that the pairing function works on two pluralities, one contributed by the distributivity key and the other contributed by a counting quantifier being distributively evaluated, it seems that Lin also has in mind some form of dynamic logic, which allows him to access the two pluralities, as well as some form of plural logic, which allows him to access the respective subparts of the two pluralities. In these respects, Lin's study can be seen as a precursor of the present study.

However, there are some pronounced differences between the Lin's analysis and the analysis proposed here. First, Lin also does not formally recognize the role of *ge* as signaling co-variation and he does not discuss how the pairing function can be constrained to model this fact. Second, although Lin also takes bare nouns to be a kind, which fails to license *ge*, the reason for that is attributed to the vague notion of intensionality. In Lin's analysis, *ge* is incompatible with intensional individuals and properties.

Lee et al. (2009a) builds on Lin (1998b) but differs from it in two important respects. First, they recognize that noun phrases with a measurement component are not the only expressions that license *ge*, so they propose a generalization of Lin's pairing function to handle this. The pairing function not only can pair a value provided by the distributivity key with a value provided by a counting quantifier, it may also pair a value provided by the distributivity key with a value provided by many other types of expressions. Second, Lee et al. (2009a) recognize a differentiation requirement of *ge* and suggests building it in as a requirement on the pairing function. Abstracting away from the differences in the level of formal explicitness, Lee et al. (2009a)'s suggestion to enrich the pairing function with a differentiation requirement is very similar to the monotonicity constraint put forward in the present study. If we take the pairing function to be our relational quantificational subordination, which allows the retrieval of a dependency using quantificational subordination, then the additional requirement for co-variation can be seen as the requirement for a monotonic mapping from the distributivity key to a relevant part in the distributivity share.

5.6 Conclusion

In this chapter, I have shown that the Mandarin distributivity marker *ge* imposes a monotonicity constraint on the functional dependencies of distributivity, just like English binominal *each*. However, it differs from binominal *each* in that its constraint is compatible with a greater variety of dependent structures. The generality of this type of monotonicity constraint can be further extended to understand the event differentiation condition of determiner and adverbial *each*.

Bibliography

Abels, K. and L. Martí (2010). A unified approach to split scope. *Natural Language Semantics 18*(4), 435–470.

Alonso Ovalle, L. and P. Menendez Benito (2003). Some epistemic indefinites. In M. Kadowaki and S. Kawahara (Eds.), *Proceedings of NELS33*, pp. 1–12.

Au Yeung, W. H. (1998). *An interface program for parameterization of classifiers in Chinese*. PhD. dissertation, Hong Kong University of Science and Technology, Hong Kong.

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

Barker, C. (2007). Parasitic scope. *Linguistics and Philosophy 30*, 407–444.

Barwise, J. (1981). Scenes and other situations. *The Journal of Philosophy 78*, 369–397.

Barwise, J. and R. Cooper (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy 4*, 159–219.

Beck, S. (2000). The semantics of *Different*: Comparison operator and relational adjective. *Linguistics and Philosophy 23*(1), 101–139.

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

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.

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. (2010). Decomposing modal quantification. *Journa of Semantics 27*, 437–527.

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.

Brasoveanu, A. and J. Dotlačil (2015). Strategies for scope taking. *Natural Language Semantics 23*, 1–19.

Brasoveanu, A. and D. Farkas (2011). How indefinites choose their scope. *Linguistics and Philosophy 34*, 1–55.

Brisson, C. (1998). *Distributivity, maximality, and floating quantifiers*. Phd. dissertation, Rutgers University, New Brunswick.

Brisson, C. (2003). Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy 26*, 129–184.

Bumford, D. (2015). Incremental quantification and the dynamics of pair-list phenomena. *Semantics & Pragmatics 8*, 1–70.

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

Cable, S. (2007). *The Grammar of Q: Q-Particles and the Nature of Wh-Fronting, as Revealed by the Wh-Questions of Tlingit*. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

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

Carlson, G. (1987). Same and different: Some consequences for syntax and semantics. *Linguistics and Philosophy 10*(4), 531–566.

Carlson, G. (1998). Thematic roles and the individuation of events. In S. Rothstein (Ed.), *Events and grammar*, pp. 35–51. Kluwer.

Carlson, G. N. (1977a). *Reference to Kinds in English*. Ph. D. thesis, University of Massachusetts, Amherst.

Carlson, G. N. (1977b). A unified analysis of the English bare plural. *Linguistics and Philosophy 1*, 413–457.

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.

Champollion, L., D. Bumford, and R. Henderson (Forthcoming). Donkeys under discussion. *Semantics & Pragmatics*.

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

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

Chen, L. (2005). *Dou* (dis)harmony in chinese. In *Proceedings of North East Linguistic Society 35*.

Cheng, L. L.-S. (1995). On dou-quantification. *Journal of East Asian Linguistics 4*(3), 197–234.

Cheng, L. L.-S. and R. Sybesma (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry 30*(4), 509–542.

Chierchia, G. (1984). *Topics in the Syntax and Semantics of Infinitives and Gerunds*. Ph. D. thesis, University of Massachusetts, Amherst.

Chierchia, G. (1992). Anaphora and dynmaic binding. *Linguistics and Philosophy 15*, 111–183.

Chierchia, G. (1993). Questions with quantifiers. *Natural Language Semantics 1*(2), 181–234.

Chierchia, G. (1995). Individual-level predicates as inherent generics. In G. N. Carlson and F. J. Pelletier (Eds.), *The Generic Book*, pp. 176–223.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics 6*(4), 339–405.

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

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese 174*, 99–149.

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

Choe, J. W. (1987b). LF movement and Pied-Piping. *Linguistic Inquiry 18*(2), 348–353.

Cohen, A. and N. Erteschik-Shir (2002). Topic, focus, and the interpretation of bare plurals. *Natural Language Semantics 10*(2), 125–165.

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

Dalrymple, M., M. Kanazawa, Y. Kim, S. A. Mchombo, and S. Peters (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 159–210.

Dawson, V. (2018). A new kind of epistemic indefinites. In U. Sauerland and S. Solt (Eds.), *Proceedings Sinn und Bedeutung 22*, Volume 1, pp. 349–366.

Dayal, V. (1996). *Locality in Wh-Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer Academic Press.

Dayal, V. (1999). Bare NP's, reference to kinds, and incorporation. In *Proceedings of Semantics and Linguistic Theory 9 (SALT-9)*.

Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy 27*(4), 393–450.

Dayal, V. (2011a). Bare noun phrases. In K. von Heusinger and P. Portner (Eds.), *Semantics: An international handbook of natural language meaning*, Volume 2, pp. 1087–1108. De Gruyter Mouton.

Dayal, V. (2011b). Hindi pseudo-incorporation. *Natural Language and Linguistic Theory 29*(1), 123–167.

Dayal, V. (2013). On the existential force of bare plurals across languages. In I. Caponigro and C. Cechetto (Eds.), *From Grammar to Meaning: The Spontaneous Logicality of Language*, pp. 49–80.

Dayal, V. (2017). *Questions*. Oxford: Oxford University Press.

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

Dekker, P. (1993). Existential disclosure. *Linguistics and Philosophy 16*(6), 561–588.

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*.

Dowty, D. (1987). Collective predicates, distributive predicates, and *all*. In A. Miller and Z.-s. Zhang (Eds.), *3rd Eastern States Conference on Linguistics (ESCOL)*, pp. 97–115.

Elworthy, D. A. H. (1995). A theory of anaphoric information. *Linguistics and Philosophy 18*(3), 297–332.

Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry 22*(1), 1–27.

Evans, G. (1977). Pronouns, quantifiers, and relative clauses. *Canadian Journal of Philosophy 7*, 467–536.

Farkas, D. (1981). Quantifier scope and syntactic islands. In R. Hendrick, C. Masek, and M. F. Miller (Eds.), *Papers from the 17th regional meeting of the Chicago Linguistic Society*, University of Chicago, pp. 59–66. Chicago Linguistics Society.

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. and H. de Swart (2003). *The semantics of incorporation: From argument structure to discourse transparency*. CSLI Publications.

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

Ferreira, M. (2005). *Event quantification and plurality*. PhD. dissertation, MIT.

Fodor, J. D. and I. Sag (1982). Referential and quantificational indefinites. *Linguistics and Philosophy 5*, 355–398.

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.

Gawron, J. M. and A. Kehler (2004). The semantics of respective readings, conjunction, and filler-gap dependencies. *Linguistics and Philosophy 27*(2), 169–207.

Geurts, B. (2000). Indefinites and choice functions. *Linguistic Inquiry 31*(4), 731–738.

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.

Grosu, A. and F. Landman (1998). Strange relatives of the third kind. *Natural Language Semantics 6*(2), 125–170.

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

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts, Amherst.

Heim, I. (1983). File Change Semantics and the the familiarity theory of definiteness. In R. Bäuerle, C. Schwarze, and A. v. Stechow (Eds.), *Meaning, Use and Interpretation of Language*, pp. 164–189. Berlin: De Gruyter.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy 13*(2), 137–178.

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

Heim, I., H. Lasnik, and R. May (1991). On “reciprocal scope". *Linguistic Inquiry 22*(1), 172–192.

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

Hintikka, J. (1973). *Logic, Language-Games and Information:Kantian Themes in the Philosophy of Logic*. Oxford: Clarendon Press.

Hintikka, J. (1986). The semantics of a certain. *Linguistic Inquiry 17*(2), 331–335.

Huang, C.-T. J. (1982). *Logical relations in Chinese and the theory of grammar*. Ph. D. thesis, Massachusetts Institute of Technology.

Huang, C.-T. J. (2002). Distributivity and reflexivity. In S.-W. Tang and C. S. Liu (Eds.), *On the formal way to Chinese Languages*.

Huang, C.-T. J. and C.-S. L. Liu (2001). Logophoricity, attitudes and Ziji at the interface. In P. Cole, G. Hermon, and C.-T. J. Huang (Eds.), *Syntax and Semantics: Long Distance Reflexives*, New York, pp. 141–195. Academic Press.

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.

Jenks, P. (2018). Articulated definiteness without articles. *Linguistic Inquiry 49*, 501–536.

Jiang, L. (2012). *Nominal arguments and language variation*. Phd, Harvard University.

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

Kamp, H. and A. Bende-Farkas (2019). Epistemic specificity. *Journal of Semantics 36*, 1–51.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.

Karttunen, L. (1976). Discourse referents. In J. D. McCawley (Ed.), *Syntax and Semantics*, Volume 7, pp. 363–385. Academic Press.

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

Keenan, E. L. and J. Stavi (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy 9*, 253–326.

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

Krahmer, E. and R. Muskens (1995). Negation and disjunction in Discourse Representation Theory. *Journal of Semantics 12*, 357–376.

Kratzer, A. (1995). Stage-level and individual-level predicates. In G. N. Carlson and F. J. Pelletier (Eds.), *The Generic Book*, pp. 125–175. Chicago: The University of Chicago Press.

Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck and L. Zaring (Eds.), *Phrase Structure and the Lexicon*, pp. 109–137. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kratzer, A. (1998). Scope or pseudoscope? Are there wide-scope indefinites? In S. Rothstein (Ed.), *Events and grammar*, pp. 163–196. Dordrecht: Kluwer Academic Publishers.

Kratzer, A. (2002). The Event Argument and the Semantics of Verbs.

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. (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.

Krifka, M. (2001). Quantifying into question acts. *Natural Language Semantics 9*(1), 1–40.

Krifka, M. and F. Modarresi (2016). Number neutrality and anaphoric update of pseudo-incorporated nominals in Persian (and weak definites in English). In *Proceedings of Semantics and Linguistic Theory 26 (SALT-26)*, pp. 874–891.

Kubota, Y. and L. Robert (2016). The syntax-semantics interface of 'respective' predication: a unified analysis in Hybrid Type-Logical Categorial Grammar. *Natural Language and Linguistic Theory 34*, 911–973.

Kuhn, J. (2014). Gather-type predicates: massiness over participants. In *Presentation at the 45th Annual Meeting of the North East Linguistic Society (NELS 45)*, http://www.jeremykuhn.net/papers/Kuhn-gather-slides.pdf.

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.

Kung, H.-I. (1993). *The mapping hypothesis and postverbal structures in Mandarin Chinese*. Ph.d., University of Wisconsin, Madison.

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

Larson, R. K. (1985). On the syntax of disjunction scope. *Natural Language and Linguistic Theory 3*(2), 217–264.

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

Lasersohn, P. (1998). Events in the semantics of collectivizing adverbials. In S. Rothstein (Ed.), *Events and Grammar*, pp. 273–292. Dordrecht: Kluwer Academic Publishers.

LaTerza, C. (2014). *Distributivity and plural anaphora*. PhD. dissertation, University of Maryland.

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.

Lee, P.-l., l. Zhang, and H. Pan (2009a). The distributive operator *ge* and some related issues. *Language and Linguistics 10*, 293–314.

Lee, P. P.-l. (2012). *Cantonese particles and affixal quantification*. Springer.

Lee, P. P.-l., L. Zhang, and H. Pan (2009b). Hanyu quancheng lianghua fuci/fengpei suanzi de gongxian he yuyi fenggong. *Hanyu Xuebao 27*, 59–70.

Lee, T. H.-T. (1986). *Studies on quantification in Chinese*. Ph.d., University of California at Los Angeles.

Lee, T. H.-T. (1994). Yueyu '*saai*' de luoji tedian [the logical properties of cantonese *saai*]. In C.-Y. Sin (Ed.), *Proceedings of the first international conference of Cantonese and other Yue dialects*, Hong Kong, pp. 131–138. Modern Educational Research Society.

Lei, K. Y. M. (2017). *The acquisition of A-quantification in Cantonese*. PhD. dissertation, Chinese University of Hong Kong, Hong Kong.

Li, H. and J. H.-K. Law (2016). At least some distributive operators aren't distributive operators. In *Proceedings of North East Linguistic Society 46*.

Lin, J.-W. (1998a). Distributivity in Chinese and its implications. *Natural Language Semantics 6*(2), 201–243.

Lin, T.-H. J. (1998b). On ge and related problems. In L. Xu (Ed.), *The referential properties of Chinese noun phrases*, pp. 209–253.

Lin, T.-H. J. (2004). Aspect, distributivity, and *wh*/qp interaction in Chinese. *Language and Linguistics 5*, 615–642.

Lin, T.-H. J. (2005). Notes on the determiner ge and related problems. In *USTWPL*, Volume 1, pp. 267–294.

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. Baeuerle, C. Schwarze, and A. v. Stechow (Eds.), *Meaning, Use and Interpretation of Language*. DeGruyter.

Link, G. (1998). *Algebraic Semantics in Language and Philosophy*. Stanford, California: CSLI Publications.

Liu, M. (2016). *Varities of alternatives*. Phd. dissertation, Rutgers University, New Brunswick.

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

Luke, K.-K. and M. L.-Y. Wong (2005). The Hong Kong Cantonese Corpus: design and uses. *Journal of Chinese Linguistics*, 312–333.

Matthewson, L. (1999). On the interpretation of wide-scope indefinites. *Natural Language Semantics 7*(1), 79–134.

May, R. (1977). *The Grammar of Quantification*. Ph. D. thesis, Massachusetts Institute of Technology.

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.

Montague, R. (1974). The proper treatment of quantification in ordinary English. In R. H. Thomason (Ed.), *Formal Philosophy: Selected Papers of Richard Montague*, pp. 247–270. New Haven: Yale University Press.

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.

Pan, H. (1998). Closeness, prominence, and binding theory. *Natural Language and Linguistic Theory 16*(4), 771–815.

Partee, B. H. (1986). Noun phrase interpretation and type-shifting principles. In D. d. J. Groenendijk and M. Stokhof (Eds.), *Studies in discourse representation theory and the theory of generalized quantifiers*, pp. 115–143.

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.

Pylkkänen, L. (2008). *Introducing Arguments*. Cambridge, Massachusetts: MIT Press.

Quine, W. V. O. (1960). *Word and object*. MIT Press.

Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy 20*(4), 335–397.

Rett, J. (2008). *Degree Modification in Natural Language*. Ph.d. dissertation, Rutgers University.

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.

Rooth, M. and B. Partee (1982). Conjunction, type ambiguity, and wide scope "or". In *Proceedings of West Coast Conference on Formal Linguistics 1*.

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.

Scha, R. (1981). Distributive, collective and cumulative quantification. In M. Stokhof and T. Janssen (Eds.), *Formal methods in the study of language*, pp. 483–512. Amsterdam: Mathematisch.

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

Schwarz, B. (2001). Two kinds of long-distance indefinites. In *Amsterdam Colloquium*.

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

Schwarzschild, R. and K. Wilkinson (2002). Quantifiers in comparatives: a semantics of degree based on intervals. *Natural Language Semantics 10*(1), 1–41.

Scontras, G. (2014). *The semantics of measurement*. Ph.d. dissertation, Harvard University.

Soh, H. L. (2005). Mandarin distributive quantifier Ge 'each', the structures of double complement constructions and the verb-preposition distinction. *Journal of East Asian Linguistics 14*(2), 155–173.

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.

Stone, M. (1992). Or and anaphora. In C. Barker and D. Dowty (Eds.), *Proceedings of Semantics and Linguistic Theory 2*, pp. 367–385.

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.

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

Szabolcsi, A. (1997). *Ways of scope taking*. Dordrecht ; Boston ; London: Kluwer Academic Publishers.

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

Tang, C.-C. J. (1989). Chinese reflexives. *Natural Language and Linguistic Theory 7*, 93–121.

Tang, S.-W. (1996). A role of lexical quantifiers. *Studies of the Linguistic Sciences 26*, 307–323.

Tomioka, S. and Y. Tsai (2005). Domain restrictions for distributive quantification in Mandarin Chinese. *Journal of East Asian Linguistics 14*(2), 89–120.

Trinh, T. (2011). Nominal reference in two classifier languages. In I. Reich, E. Horch, and D. Pauly (Eds.), *Proceedings Sinn und Bedeutung 15*, pp. 629–644.

Tsai, Y. (2009). *Aspects of distributivity in Mandarin Chinese*. Ph. D. thesis, University of Delaware.

Tunstall, S. (1998). *The Interpretation of Quantifiers: Semantics and Processing*. Ph. D. thesis, University of Massachusetts at Amherst, Amherst, Massachusetts.

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

Vendler, Z. (1962). *Every* and *each*, *any* and *all*. *Mind 71*, 145–160.

von Heusinger, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 245–274.

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

Winter, Y. (2002). Functional readings and wide-scope indefinites. In B. Jackson (Ed.), *Proceedings of SALT XII*, Cornell University, Ithaca, NY, pp. 306–321. CLC Publications.

Xiang, M. (2008). Plurality, maximality and scalar inferences: A case study of Mandarin *Dou*. *Journal of East Asian Linguistics 17*(3), 227–245.

Xiang, Y. (2016). *Interpreting Questions with Non-exhaustive Answers*. Ph. D. thesis, Harvard University, Cambridge, Massachusetts.

Yang, R. (2001). *Common nouns, classifiers, and quantification in Chinese*. Ph. D. thesis, Rutgers University.

Yanovich, I. (2005). Choice-functional series of indefinite pronouns and Hamblin semantics. In E. Georgala and J. Howell (Eds.), *Proceedings of SALT XV*, Cornell University, Ithaca, NY, pp. 309–326. CLC Publications.

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.