

DISTRIBUTIVITY, LEXICAL SEMANTICS, AND WORLD KNOWLEDGE

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF LINGUISTICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Lelia Montague Glass

June 2018

© 2018 by Lelia Montague Glass. All Rights Reserved.

Re-distributed by Stanford University under license with the author.



This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.

<http://creativecommons.org/licenses/by-nc/3.0/us/>

This dissertation is online at: <http://purl.stanford.edu/st374mm5103>

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Beth Levin, Co-Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Christopher Potts, Co-Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Cleo Condoravdi

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Daniel Lassiter

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

A predicate is understood distributively if it is inferred to be individually true of each member of a plural subject, nondistributively if not. *Alice and Bob smiled* conveys that Alice smiled and Bob smiled (distributive); *Alice and Bob met* conveys that they met jointly (nondistributive); *Alice and Bob opened the window* can describe a situation in which they each did so (distributive), or one in which they did so only jointly (nondistributive).

These facts raise a compositional semantics question and a lexical semantics question. The compositional semantics question has been discussed widely: how should these sentences be represented semantically? To what extent should such representations capture inferences about distributivity? The lexical semantics question has received less attention: which predicates are understood in which ways? Certainly these inferences are grounded in the events described by these predicates (*smile* is distributive because people have their own faces and can only *smile* individually); but which further predicates behave like *smile*, like *meet*, or like *open the window*, and why?

To make progress on the lexical semantics question, this dissertation presents the Distributivity Ratings Dataset, over 2300 verb phrases (built from the verbs of [Levin 1993](#)) rated for their distributivity potential by online annotators. This dataset provides evidence consistent with a series of far-reaching hypotheses: that predicates describing the action of an individual body or mind (*smile*, *faint*, *swallow a pill*) are distributive given that individuals have their own bodies and minds; that predicates describing inherently multilateral actions (*meet*, *gather*) are nondistributive given that individuals such as *Alice* cannot carry out these actions unilaterally; that causative predicates (*open a window*, describing an action where the subject causes the object to change) can (but need not) be nondistributive given that multiple individuals' actions may be jointly but not individually sufficient to cause a result; and finally, that predicates with incremental objects (objects whose parts correspond to the parts of the event described by the predicate, as in *eat the pizza*) can also be nondistributive, given that each member of a plural subject might carry out the verb event on a different portion of the object, only jointly adding up to the whole.

Turning from verb phrases to adjectives, the dissertation draws on tools from measurement the-

ory to argue that a gradable adjective’s potential for distributivity depends on the nature of the scale associated with it (assuming that gradable adjectives relate individuals to ‘degrees’ along a scale). A predicative adjective can be understood nondistributively (as when *the boxes are heavy* conveys that the boxes are jointly but not individually heavy) if the scale associated with the adjective behaves ‘positively’ with respect to concatenation: if the weight of two boxes together exceeds the weight of each one. That way, the contextual standard for what counts as *heavy* can be set in such a way that two boxes together exceed it, while each box individually falls short of it — nondistributive, because *heavy* is true of the two boxes together, but not of each one alone. Other adjectives are not associated with scales that behave in this way, explaining why they are only understood distributively: *the boxes are new* conveys that each box is new (distributive), not that they are new jointly, because two boxes together are no newer than each one. In sum, this dissertation puts forward a series of large-scale generalizations about how the distributivity potential of various verbal and adjectival predicates is derived from the nature of the events and properties that they describe.

Turning to the compositional semantics question, the dissertation advocates for an underspecified semantic representation in which a predicate is true of each cell of a contextually supplied cover (set of subparts) of its plural subject. All inferences about distributivity are framed as inferences about which cover(s) to entertain, given what is known about the event or property described by the predicate and how the members of the subject can participate in it. This semantic analysis does not explain anything on its own, but becomes explanatory when combined with a predictive analysis of which predicates can be understood in which ways. In this way, the compositional semantics question and the lexical semantics question are framed as complements to one another: an underspecified compositional representation is supplemented with an articulated theory of how a predicate’s distributivity potential depends on the nature of the event or state it describes.

While distributivity has traditionally been studied as a topic for compositional semantics, it is defined by the observation that different predicates (*smile, meet, open the window; heavy, new*) act differently from one another, making it a lexical semantics topic from the start. This dissertation aims to illuminate it by treating it as one.

Acknowledgements

I am grateful above all to my committee: Beth Levin and Christopher Potts (my co-chairs); Cleo Condoravdi, and Daniel Lassiter.

I came to Stanford because I was inspired by Beth's lexical semantics course at the 2011 LSA Institute, and Chris's work expanding the phenomena studied in semantics. Although Beth and Chris work in different areas and have never before co-chaired a dissertation, they are deeply similar in that they contribute empirical rigor and transcendent insights in all their work, and they make an otherwise abstract topic tractable by tying it to independently motivated human reasoning. Beth guided my development as a lexical semanticist, Chris helped me grow as a pragmaticist, and both of them taught me to collect data, all of which fundamentally shaped this dissertation. Methodologically, I was particularly influenced by a project on compounds that I did with Beth Levin and Dan Jurafsky. Beth also has my deepest gratitude for her guidance in my career more generally.

I am grateful for Cleo not just for helping me better engage with the distributivity literature, but especially for her patience and optimism. She saw the potential of this topic before anyone else could (including me); without her, I might have abandoned it before giving it time to blossom. I thank Dan L. for pushing me to do more rigorous statistical analyses, for helping me frame my contribution, and for modeling a general approach to semantics and pragmatics which I have been inspired by. Thanks also to my external chair, Mark Crimmins.

My dissertation has also benefitted greatly from conversations I've had with Dylan Bumford, Heather Burnett, Lucas Champollion, James N. Collins, Jeremy Kuhn, Louise McNally, Jessica Rett, Roger Schwarzschild, Gregory Scontras, Hanna de Vries, Alexander Williams, and Yoad Winter. Lucas Champollion deserves particular thanks for generously taking the time to send me detailed comments on several materials.

This work was presented at the 43rd Berkeley Linguistics Society conference; the Linguistic Society of America meeting in Salt Lake City; the Linguistic Evidence conference in Tübingen; the CNRS Journées (Co-)Distributivité in Paris; the University of Utrecht; and as an invited talk at various other places. I am grateful for constructive comments at all these venues.

For help collecting data, I thank Nanjiang Jiang, who worked with me as a summer intern in 2017; and the people of Amazon’s Mechanical Turk for their care and effort. More generally, I thank the creators of open-source software used in this dissertation (L^AT_EX, Python, and R), as well as researchers who make their work easily available.

This project began while I was an intern in the Natural Language Processing and Artificial Intelligence Laboratory at Nuance Communications, when Kathleen Dahlgren and Karen Wallace asked me to code a set of verbs for their distributivity potential, setting me on my current path. Without them, I don’t think I would have ever worked on distributivity.

Financially, I am grateful to the Stanford Department of Linguistics, the Stanford Vice Provost for Graduate Education, the American Council of Learned Societies (ACLS), and the Phi Beta Kappa Northern California Association for supporting this work.

I became a linguist because of my experience as an undergraduate at the University of Chicago, where I was inspired (and welcomed, before I even knew anything) by Karlos Arregi, Itamar Francez, Anastasia Giannakidou, Chris Kennedy, Jason Merchant, Malte Willer, and Ming Xiang. I also had the privilege of learning from some semanticists who were then graduate students in that department — Rebekah Baglini, Andrea Beltrama, M. Ryan Bochnak, and Peter Klecha.

At Stanford, I have been lucky to be part of an energetic semantics / pragmatics community, for which I especially thank Eric K. Acton, Rebekah Baglini, Samuel R. Bowman, Dylan Bumford (who visited for awhile), Reuben Cohn-Gordon, James N. Collins, Cleo Condoravdi, Phil Crone, Judith Degen, Alex Djalali, Masoud Jasbi, Sunwoo Jeong, Lauri Karttunen, Sara Kessler, Bonnie Krejci, Emily Lake, Daniel Lassiter, Sven Lauer, Beth Levin, Prerna Nadathur, Stanley Peters, Christopher Potts, and Annie Zaenen.

It has also been a privilege to serve as a mentor for the EDGE (Enhancing Diversity in Graduate Education) program at Stanford. I’m grateful to Chantal Gratton (my mentee), and to Solomon Hughes and Chris Gonzalez Clarke for creating the amazing cross-disciplinary EDGE community.

One main reason to do linguistics is the people it attracts. I’ve enjoyed the friendship and supportive spirit of my cohort: Philip Crone, Timothy Dozat, Katherine Hilton, Masoud Jasbi,

Sharese King, Bonnie Krejci, and Teresa Pratt. It looks like one hundred percent of us will complete the Ph.D.! Other great linguist friends include Sam Bowman, James N. Collins, Sunwoo Jeong, Sara Kessler, Ed King, Judit (Judy) Kroo, Emily Lake, Daisy Leigh, Kate Lynn Lindsey, Prerna Nadathur, and Simon Todd (who also has helped me quite a lot with statistics — thank you, Simon!).

Outside of linguistics, I owe a great deal to my lifelong friends Leslie Adkins, Emma Carlin, Katherine Gao, Paige Gresty, Anna Schleusener, Caroline Wooten, and Amanda Yeager. I will always be grateful to my grandparents, Gayle Schoenfeldt Glass and Carter Monroe Glass, for supporting my education at the University of Chicago, which truly changed my life. My fiancé Mark Menzies has been there from my first term papers through my dissertation defense, and has always believed in me. Finally, I thank my parents, Dale Soutter Glass and Carter Martin Glass (Dad, we miss you!), for supporting my dream of becoming a linguist. This dissertation is for them.

Contents

1	Introduction	1
1.1	What is distributivity?	1
1.2	Plan of attack	3
1.2.1	Main questions	3
1.2.2	Preview of claims	4
1.2.3	Guiding principles	5
1.2.4	Distinguishing linguistic and non-linguistic knowledge	6
1.3	Complications	9
1.3.1	Types of subjects	9
1.3.2	Arguments other than the subject	10
1.3.3	The effect of the object of a transitive verb	11
1.3.4	What's possible versus what's preferred	14
1.4	Outline of the dissertation	15
2	'Collective' vs. 'cumulative'	19
2.1	Introduction	19
2.2	Should 'collective' be separate from 'cumulative'?	22
2.2.1	For and against defining collectivity positively	22
2.2.2	For and against a collective / cumulative distinction	27
2.3	Cumulativity of verbs and thematic roles	34

2.4	Evidence from predicates with incremental objects	38
2.5	Chapter summary	43
3	Semantic representation	44
3.1	Introduction	44
3.2	Data to capture	45
3.3	A cover analysis	48
3.3.1	Schwarzschild’s formulation	48
3.3.2	Analysis advocated here	52
3.3.3	Capturing the ‘collective’ / ‘cumulative’ data on the proposed analysis . . .	56
3.3.4	Discussion	59
3.4	Alternative analyses from the literature	60
3.4.1	One source: an operator	60
3.4.2	One source: meaning postulates	62
3.4.3	Two sources: meaning postulates and an operator	63
3.4.4	Discussion	68
3.5	Chapter summary	68
4	Verb phrases	69
4.1	Introduction	69
4.1.1	Literature motivating the current study	70
4.1.2	Where the current work fits in	72
4.2	Distributivity Ratings Dataset	74
4.2.1	Choosing objects for transitives	76
4.2.2	Study design	79
4.2.3	Results	85
4.3	Motivating and testing hypotheses	88
4.3.1	Full models including all predictors	88

4.3.2	Transitive / intransitive asymmetry	98
4.3.3	Body / mind predicates	101
4.3.4	Multilateral predicates	106
4.3.5	Causatives	110
4.3.6	Predicates with incremental objects	116
4.3.7	Discussion	123
4.4	Chapter summary	125
5	Adjectives	126
5.1	Introduction	126
5.2	Literature on the distributivity of adjectives	129
5.2.1	A pragmatic explanation for <i>heavy</i> versus <i>tall</i>	129
5.2.2	Open questions	132
5.3	Background on gradable adjectives and measurement theory	134
5.4	Explaining the distributivity potential of adjectives	139
5.5	Chapter summary	149
6	Conclusion	151
6.1	Summary	151
6.2	Open questions	152
6.3	Zooming out	154

List of Tables

1.1	(In)definiteness and (non)repeatability interact to constrain a predicate’s potential for a distributive understanding.	14
4.1	Average ratings for both ‘each’ and ‘together’ questions for each predicate.	85
4.2	Each participant’s ratings for both the ‘each’ and ‘together’ questions for each predicate they encountered.	86
4.3	Number and percentage of responses in each of the five response categories for both the ‘each’ and ‘together’ questions.	86
4.4	Number of predicates in each category, and overlap between the categories.	90
4.5	Model estimates for the maximal ‘each’ model (allowing all interactions that make sense), with random intercepts for both participants and predicates.	94
4.6	Model estimates for the most parsimonious and predictive ‘each’ model according to the Akaike Information Criterion — allowing only one interaction, with random intercepts for both participants and predicates. The statistics reported below come from this model.	95
4.7	Model estimates for the maximal ‘together’ model (allowing all interactions that make sense), with random intercepts for both participants and predicates.	96

4.8	Model estimates for the most parsimonious and predictive ‘together’ model according to the Akaike Information Criterion — with random intercepts for both participants and predicates, but no interactions. The statistics reported below come from this model.	97
5.1	Distributivity potential of different types of adjectives.	128
5.2	Proposed explanation for why some adjectives are distributive, some can be understood in both ways, and some are ‘stubbornly distributive’.	149

List of Figures

4.1	Screenshot of the instructions page.	83
4.2	Screenshot of an item from the experiment.	84
4.3	Verb phrases built from transitive verbs have systematically lower ‘each’ ratings, and systematically higher ‘together’ ratings, compared to intransitives.	100
4.4	Body / mind intransitives have systematically higher ‘each’ ratings, and systematically lower ‘together’ ratings, than other intransitives. In the same way, body / mind transitives have systematically higher ‘each’ ratings, and systematically lower ‘together’ ratings, than other transitives.	105
4.5	Multilateral verbs (all intransitive) have lower ‘each’ ratings and higher ‘together’ ratings than other intransitives.	109
4.6	Causatives (all transitive) have lower ‘each’ ratings and higher ‘together’ ratings than other transitives.	115
4.7	Predicates with objects that can be construed as incremental (all built from transitive verbs) have lower ‘each’ ratings and higher ‘together’ ratings than other transitives.	124
5.1	Distributive and nondistributive understandings of <i>heavy</i>	140
5.2	The boxes together qualify as <i>tall</i> , but not individually.	143
5.3	<i>Light</i> is true of things lighter than the contextual standard θ (here, 7lbs).	144
5.4	The cake and the fudge together are no <i>warmer</i> than they are individually.	145

6.1 Verb phrases built from transitive verbs have systematically lower ‘each’ ratings,
and systematically higher ‘together’ ratings, compared to intransitives. 158

‘I have tried to dispel the misconception [...] that all the interesting and important problems of natural language semantics have to do with so-called logical words [...] rather than word-semantics, as well as with the more basic misconception that it is possible even to separate these two kinds of problems.’

— Dowty 1979 (Foreword)

Comments and questions are most welcome.

glasslelia@gmail.com

Chapter 1

Introduction

1.1 What is distributivity?

The goal of semantics and pragmatics is to understand the inferences that people draw from (uses of) sentences. This dissertation zooms in on a particular class of inferences: those drawn from a sentence with a plural subject, such as *the children*, about how each member of the subject (each child) participates in the predicate of the sentence.¹

In (1), we infer that the children *each* smiled. This way of understanding *smile* is described as ‘distributive’, because *smile* ‘distributes’ to — is individually true of — each member of the subject *the children*. (It does not matter whether the children were interacting ‘together’ when they smiled; all that matters for (1) to be understood distributively is that *smile* is true of each child.)

(1) The children smiled.

- a. ✓**Distributive:** The children *each* smiled.
- b. ✗**Nondistributive:** The children smiled jointly without each individually doing so.

¹Key references include Link 1983, Dowty 1987, Roberts 1987, Landman 1989a, Lasersohn 1995, Schwarzschild 1996, Winter 1997, Winter 2000, Landman 2000, Champollion 2010, de Vries 2015, and Champollion 2017; while Lasersohn 2011, Nouwen 2015, Winter & Scha 2015, and Champollion to appear provide introductory overviews.

In contrast, we do not infer from (2) that the children *each* met — unless we reinterpret *meet* to have an implicit object (met *with someone*). Instead, we infer that the children met jointly, not individually. This understanding can be described as ‘nondistributive’, in that *meet* does not ‘distribute’ (apply separately) to each child, but rather seems to apply to *the children* as a whole.

- (2) The children met.
- a. **✗Distributive:** The children *each* met.
 - b. **✓Nondistributive:** The children met jointly without each individually doing so.

In the literature, nondistributive understandings are also called ‘collective’, a term which is sometimes associated with inferences about collaboration and joint responsibility (Landman 2000, Champollion 2010). In Chapter 2, I take on the question of what distributivity should be contrasted with. For now, I simply call (2b) ‘nondistributive’. Such an understanding is practically unimaginable for *smile* (1b), while for *meet*, it is the natural one.

While *smile* is understood distributively, and *meet* is understood nondistributively, other predicates can be understood in both ways. (3) could describe a situation in which each child opened the window, one after another (3a); or could describe a situation in which the children opened the window jointly (3b), for example by pushing on it all at once.

- (3) The children opened the window.
- a. **✓Distributive:** The children *each* opened the window.
 - b. **✓Nondistributive:** The children opened the window jointly without each individually doing so.

As for terminology: predicates which can be understood both distributively and nondistributively, like *open the window*, are sometimes called ‘mixed’ predicates (Link 1983, Dowty 1987), based on the idea that they can have in their extension both atomic individuals such as *Alice*, and multi-part groups or pluralities such as *the children*. The literature also refers to distributive and col-

lective ‘readings’ of such predicates, on the assumption that each of these correspond to a distinct logical representation of the sentence. In addition to ‘mixed’ predicates like *open the window*, predicates like *smile* are often called ‘distributive’ predicates, while those like *meet* are called ‘collective’ predicates.

In this dissertation, I want to expose the inferential process underlying this classification. While the term ‘reading’ connotes a semantic ambiguity, I argue (Chapter 3) that the different ways of understanding a predicate such as *open the window* are not necessarily to be derived from distinct semantic representations (thus, I agree with [Schwarzschild 1994](#), [Verkuyl & van der Does 1996](#), [Schwarzschild 1996](#), [Moltmann 1997](#), [Nouwen 2015](#)). Using terminology that I see as more theoretically neutral, I refer to the ‘understandings’ ([Lasersohn 1990b](#): 8, [Nouwen 2015](#)) available to various predicates: *open the window* can be understood both distributively and nondistributively, *smile* is understood distributively, and *meet* is understood nondistributively.

1.2 Plan of attack

This picture raises two main theoretical questions, a compositional semantics question which has been discussed widely, and a lexical semantics question which has received less attention.

1.2.1 Main questions

First, the compositional semantics question (see, among others, [Link 1983](#), [Dowty 1987](#), [Roberts 1987](#), [Landman 1989a](#), [Lasersohn 1995](#), [Schwarzschild 1996](#), [Winter 1997](#), [Landman 2000](#), [Winter 2002](#), [Champollion 2010](#), [de Vries 2015](#), [Champollion 2017](#)): how should sentences like (1)–(3) be represented semantically? To what extent should the inferences we draw from these sentences be read off from their semantic representation? In particular, how do we capture the two distinct understandings available to (3) — in terms of a semantic ambiguity, or an underspecified meaning compatible with multiple different situations? (The term ‘understanding’ is chosen to avoid presupposing an ambiguity, but many authors do posit one.) And if (3) is ambiguous, why are (1) and (2)

unambiguous?

Second, the lexical semantics question: which other predicates behave like *smile*, like *meet*, or like *open the window* — and why? Uncontroversially, the inferences drawn from these predicates are shaped by our knowledge about how events of smiling, meeting, and window-opening take place in the world (Dowty 1987, Roberts 1987, Winter 2000, Champollion 2010, de Vries 2015; to my knowledge, no author would outright disagree). *Smile* is understood distributively in (1) because smiling is an action of the face; people have their own faces, so they cannot *smile* jointly without also doing so individually. *Meet* is understood nondistributively in (2) because it describes an inherently social action that cannot be undertaken unilaterally. *Open the window* can be understood in both ways in (3) because it can be carried out individually or jointly. But while researchers agree that world knowledge is fundamental to the inferences in (1)–(3), it is still an open question which other predicates behave like *smile*, *meet*, or *open the window*, and why.

1.2.2 Preview of claims

In this dissertation, both questions are tackled together. Concerning the compositional semantics question, I present an underspecified semantics which simply requires the predicate to be individually true of each cell of a pragmatically supplied cover — a set of subparts — of the subject (Higginbotham 1981, Gillon 1987, Schwarzschild 1996; Chapter 3). This semantics handles *smile*, *meet*, and *open the window* in a uniform way. Different inferences are drawn from these different predicates because we entertain different covers for each one. Given that people can only *smile* individually, the only sensible cover for *smile* places each child in their own cell (distributive). Given that people can only *meet* multilaterally, the only sensible cover for *meet* places multiple children in the same cell (nondistributive). Given that people can *open the window* individually or jointly, we entertain a cover placing each child in their own cell (distributive), as well as one placing all of the children in the same cell (nondistributive). Several alternative analyses (reviewed in Chapter 3) also largely capture the facts, so the cover analysis is chosen only because I see it as the most straightforward.

On its own, the proposed analysis (like its alternatives) does not make any predictions about which predicates are understood in which ways. That gap is filled by addressing the lexical semantics question: by pinpointing the aspects of world knowledge about various events and properties that shapes the distributivity potential of the predicates describing them. Chapter 4 motivates and tests hypothesized patterns within a dataset of verb phrases rated for their distributivity potential by online annotators; Chapter 5 analyzes the distributivity potential of gradable adjectives in terms of the structure of the scales associated with them.

In other words, the two questions — the compositional semantics question of how to represent distributivity and nondistributivity semantically, and the lexical semantics question of which predicates are understood in which ways — are framed as complements to one another. The compositional semantics question is answered in a way that leaves much of the work to world knowledge, and the lexical semantics question is answered in a way that aims to make the call to world knowledge explanatory.

1.2.3 Guiding principles

This dissertation is guided by the idea that it is more parsimonious, when possible, to explain a given phenomenon in terms of independently motivated, general reasoning than in terms of (silent) linguistic structure (Bar-Hillel 1971, Grice 1989), particularly because one would need such reasoning anyway in order to posit the correct silent material (e.g., Potts et al. 2016). This principle is what leads me to be skeptical of various purported semantic ambiguities in the literature — between so-called ‘collective’ and ‘cumulative’ readings of predicates (Chapter 2); between the presence or absence of the ‘group’-forming operator (Chapter 2); between the presence or absence of the *D* (distributivity) operator (Chapter 3) — when the data motivating these distinctions can be explained in other terms.

Of course, ‘general reasoning’ is only explanatory if we explain it, so this dissertation is also guided by the goal of taking on this challenge. Any time an inferential or grammatical phenomenon depends on world knowledge, the next step is to explain what world knowledge matters and why.

This principle is what leads me to investigate the factors shaping the potential for distributivity of various types of predicates.

This dissertation also takes the view that a widespread phenomenon such as distributivity should be studied by considering a wealth of data. It is valuable to analyze clean, prototypical examples such as (1)–(3), but it is equally important to test the theory against a large quantity of additional predicates.

Finally, while distributivity has generally been studied within the tradition of compositional semantics, this dissertation is guided by the idea that it must simultaneously be understood as an endeavor for lexical semantics (see the Foreword to Dowty 1979 for discussion of why compositional semantics and lexical semantics should be undertaken together). The defining data (1)–(3) illustrate that different predicates act differently with respect to distributivity; and any time that different words behave differently, grammatically or inferentially, we need lexical semantics to tell us why.

1.2.4 Distinguishing linguistic and non-linguistic knowledge

Before proceeding, it is worth briefly reviewing the terms ‘semantics’, ‘pragmatics’, ‘lexical semantics’, and ‘world knowledge’.

For background, there is a longstanding debate about how to draw a line between semantics (often defined as the literal, entailed, context-independent meaning of a sentence) and pragmatics (often defined as the non-literal, non-entailed, context-sensitive inferences drawn about why a speaker decided to utter that sentence).²

Some inferences can be relatively easily classified as semantic or pragmatic; for a sentence like *I'm tired*, many researchers would agree that semantics should deliver the inference that the speaker is tired (which, if denied, creates a contradiction), while pragmatics should handle the inference that the speaker does not want to go out (which can be denied without contradiction). Other inferences are harder to classify; there is a debate about whether certain implicatures arise grammatically or

²See, among very many others, Morris 1938, Lewis 1969, Kaplan 1977, Grice 1989, Levinson 1983, Levinson 2000, Taylor 2001, Cappelen & Lepore 2005, Szabó 2008, Carston 2008, Borg 2012, McNally 2013.

conversationally (e.g., Chierchia 2004, Russell 2006, Potts et al. 2016). As a different type of example, (4a) conveys (4b); but is that a semantic entailment, a fact about geography, or both?

- (4) a. I went to Hong Kong.
b. I went to Asia.

More generally, it is not always clear how the semantics and / or pragmatics should handle information that may be considered ‘world knowledge’ (Gamut 1991: Chapter 6). If semantics aims to capture entailment relations between sentences, then perhaps the inference from (4a) to (4b) should be considered semantic, because it behaves like an entailment (it cannot be cancelled). But if semantics is meant to capture speakers’ knowledge of a particular language, then perhaps this inference is not semantic, because a geographically ignorant English speaker could understand (4a) and (4b) without knowing that Hong Kong is in Asia. If the inference from (4a) to (4b) is not semantic, then perhaps it is pragmatic — not in the sense of a conversational implicature about why a speaker chose to say one thing over another (Grice 1989), but on a broader understanding of ‘pragmatics’ as background assumptions and general reasoning (e.g., Langacker 1987, Levinson 2000, Taylor 2001). Another ‘pragmatic’ inference in this sense would be that if (4a) is uttered in the United States, hearers infer that the speaker probably travelled by airplane.

In addition to debating how semantics can be separated from pragmatics, researchers also debate whether there is a distinction between ‘lexical semantics’ (knowledge of word meaning) and ‘world knowledge’ (knowledge of the world). Sometimes described as the ‘dictionary / encyclopedia debate’, this issue surfaces in the literatures on lexical semantics (e.g., Fillmore 1969, Pustejovsky 1995, Neeleman & Van de Koot 2012), compositional semantics (e.g., Gamut 1991: 170-173), philosophy of language (e.g., Katz & Fodor 1963, Searle 1978³), computational linguistics (e.g., Hobbs 1987), Distributed Morphology (e.g., Harley & Noyer 2000), syntax (e.g., Chomsky 1973), neurolinguistics (e.g., Hagoort et al. 2004), and elsewhere. For a thorough review, I refer to Peeters

³Illustrating the pervasive effect of world knowledge on linguistic communication, Searle 1978: 216 memorably points out that when we order a burger, we do not bother to specify that it should be a few inches in diameter and served on a plate, rather than a mile wide and encased in plastic.

2000 and references therein.

The ‘dictionary / encyclopedia debate’ has consequences for the distinction between semantics and pragmatics. The meanings of words of course contribute to the meaning of a whole sentence. If there is no dividing line between word meaning and world knowledge, then the meaning of a sentence would comprise an unbounded amount of information about each constituent word (and equivalently, the things described by each word), making it very difficult to separate linguistic and non-linguistic knowledge or reasoning, and thus to separate semantics and pragmatics. So if one does want to distinguish between semantics and pragmatics, it seems that one must draw some distinction between lexical knowledge and world knowledge.

One way to draw such a distinction is to separate arbitrary, language-specific facts from non-arbitrary, language-independent ones. The idea is that the lexicon is at least partially arbitrary and idiosyncratic (e.g., Bloomfield 1933: 274, Chomsky & Halle 1968: 12, Lieber 1980: 63); at least the mapping between form and meaning is (de Saussure 1916), as evidenced by the fact that different languages use different form / meaning mappings (*dog* refers to dogs in English; *chien* does so in French). In contrast, the world (and our ‘encyclopedic’ knowledge of it) may be systematic. The lexical fact that *ancestor* refers to ancestors is an arbitrary convention of English, but the encyclopedic fact that a father’s ancestors are also his biological son’s ancestors (Schwarzschild 1996) is non-arbitrarily explained by the biology of ancestry (5a)–(5b), and of course does not depend on what language is spoken. Thus, even though the inference from (5a) to (5b) is an entailment, Schwarzschild argues that it falls outside the jurisdiction of semantic theory.

- (5) a. Bill’s biological father has red-headed ancestors. adapted Schwarzschild 1996: 187
 b. Bill has red-headed ancestors.

For the purposes of this dissertation, I distinguish between linguistic and non-linguistic knowledge, between semantics and pragmatics, and between lexical knowledge and world knowledge. (Even if these distinctions are fuzzy and only exist in our minds, I still argue that they prove useful.) In terms of linguistic knowledge: semantics characterizes the literal meaning (truth conditions) of

sentences and the way they are assembled compositionally. As for non-linguistic knowledge: world knowledge is what we know (or believe) about the world; pragmatic reasoning is what we believe about our interlocutors (their beliefs and intentions), and more generally what we infer from a sentence above and beyond its literal meaning. ‘World knowledge’ and ‘pragmatic reasoning’ thus blend together. Straddling the division between linguistic and non-linguistic knowledge, lexical semantics characterizes the mapping between words and the things they refer to, and seeks to classify words based on the features of their referents that shape their grammatical or inferential behavior.

Returning to distributivity, these assumptions raise a question of whether (or to what extent) inferences about distributivity should be explained linguistically versus non-linguistically. A linguistic explanation of such inferences might draw on the properties of particular (language-specific) words such as English *smile*, or might posit a covert quantifier (like a silent version of *each*; see Chapter 3) in the logical representation of a sentence. A non-linguistic explanation would focus on the nature of the (language-independent) events described by particular predicates and the way the members of the sentential subject can participate in those events. Researchers already agree that at least some inferences about distributivity are grounded in non-linguistic facts (again, no one would dispute that *smile* is distributive not because of any arbitrary feature of the English word, but because people have their own faces and can only *smile* individually). This dissertation aims to generalize that type of explanation.

1.3 Complications

Here, I acknowledge complications to the data in (1)–(3), some of which I set aside.

1.3.1 Types of subjects

Fundamentally, inferences about distributivity are inferences about how the predicate of the sentence applies to the parts of the subject. Defined in this general way, we expect to find distributivity with all sorts of subjects that have parts: various types of plurals such as *the children*, *some children*, *three*

children, and so on; conjunctions such as *Alice and Bob*; and group nouns such as *the committee* (Barker 1992, de Vries 2015).

When the subject is a plural definite such as *the children*, we encounter the issue of non-maximality observed by Dowty 1987, Brisson 1998, and others: *the children smiled* may be used if only some or most of the children actually smiled.⁴ In contrast, when the subject is a conjunction such as *Alice and Bob*, nonmaximality does not arise; both individuals are inferred to have participated in the event described by the predicate, because otherwise there would be no reason to mention each one (Landman 1989b). For this reason, I use conjoined names when the nonmaximality issue would confound the data. (Winter 2001a warns against using conventionalized conjunctions such as *Simon and Garfunkel*, a music group, but none of the conjoined names I use have this status.) I do not deal with group nouns such as *committee* (see de Vries 2015 for discussion).

1.3.2 Arguments other than the subject

So far, I have defined distributivity as an inference about how the predicate of a sentence applies to members of the subject. But the same concept can be applied to parts of a sentence other than the subject (Dowty 1987, Lasersohn 1993, Champollion to appear). For example, the verb *read* is understood distributively on its object argument, in that if multiple proposals are read, they each are (6). In contrast, *summarize* can arguably be understood nondistributively as well as distributively on its object (7) (Dowty 1987), in that multiple proposals could be summarized into a single document, without each being summarized individually.

- (6) Alice read the proposals.
- a. ✓**Object-distributive:** Alice read each proposal.
 - b. ✗**Object-nondistributive:** Alice read the proposals (together), but did not read each one.

⁴In fact, as noted by Dowty 1987 and Yoon 1996, non-maximality interacts in interesting ways with lexical semantics and pragmatics: Dowty notes that *the reporters asked questions* may convey that only some reporters did so, while *the reporters were silent* conveys that nearly all of them were. See Malamud 2012, Križ 2016, and Champollion et al. to appear for discussion.

- (7) Alice summarized the proposals. adapted Dowty 1987: 106
- a. ✓**Object-distributive:** Alice summarized each proposal (individually).
 - b. ✓**Object-nondistributive:** Alice summarized the proposals (together into a single document), but did not summarize each one (individually).

For simplicity, this dissertation focuses on distributivity involving the subject position, and mainly uses example sentences with singular objects rather than plural ones. (Sentences with multiple plurals — in subject and object position — are discussed in Chapter 2.)

1.3.3 The effect of the object of a transitive verb

This dissertation focuses on the distributivity potential of predicates built from verbs, such as *smile*, *meet*, and *open the window* (turning to adjectives in Chapter 5). Of course, for a predicate built from a transitive verb such as *open*, its distributivity potential is shaped not just by the verb, but also by its object — both its referent and its grammatical properties.

If the object of a verb such as *open* refers to a body part (*open an eye*), then the resulting predicate may only have a distributive understanding, since people have their own eyes. If its object is quite small (*open a soda*), it may also favor a distributive understanding, given that sodas are easily opened and often consumed by individuals. If the object is quite large (*open a vault*), that may favor a nondistributive understanding, since it might be difficult to open a vault alone. (These issues resurface in Chapter 4, where objects are chosen for the transitive verbs tested in an online ratings study.)

When the object of a predicate is a singular count noun (plural objects are discussed in Chapter 2), it also matters whether that object is definite or indefinite; and further, whether the action described by the verb can be repeated on the same object (*Champollion to appear*).

If the object of the verb is *definite* and the action described by the verb *can* be plausibly repeated on the same object (such as *open the window*, where the same window can be opened multiple times), then a distributive understanding is sensible, as in (8).

- (8) The children opened the window. (= (3))
- a. ✓**Distributive:** The children *each* opened the window.
 - b. ✓**Nondistributive:** The children opened the window jointly without each individually doing so.

If a predicate's object is *definite* and the action *cannot* plausibly be repeated on the same object (*break the vase*), then the predicate does not make sense distributively: since the same vase cannot generally be broken more than once, the distributive understanding (9a) is implausible.

- (9) The children broke the vase.
- a. ✗**Distributive:** The children *each* broke the vase.
 - b. ✓**Nondistributive:** The children broke the vase jointly without each individually doing so.

If a predicate's object is *indefinite* and the action described by the verb *cannot* be repeated on the same object (*break a vase*), then the only sensible distributive understanding is one where the indefinite 'covaries' with each member of the subject — in (10b), each child breaks a different vase (what Dotlačil 2010 calls a distributive understanding 'with covariation'.) The distributive understanding *without* covariation, (10a), is implausible given that the same vase cannot generally be broken multiple times.

- (10) The children broke a vase.
- a. ✗**Distributive without covariation:** There is one vase; each child broke it.
 - b. ✓**Distributive with covariation:** Each child broke a different vase.
 - c. ✓**Nondistributive:** There is one vase; the children jointly broke it.

If the object is *indefinite* and the action *can* be repeated on the same object (*open a window*), then two distributive understandings are available, one with covariation (11a) and one without (11b)

(Winter 2000).

- (11) The children opened a window.
- a. ✓**Distributive without covariation:** There is one window; each child opened it.
 - b. ✓**Distributive with covariation:** Each child opened a different window.
 - c. ✓**Nondistributive:** There is one window; the children opened it jointly but not individually.

Beyond definiteness and repeatability, the nature of the action described by the verb determines whether the predicate also can be understood *nondistributively*. Whether the object of *see* is definite or indefinite, it is only understood distributively (12)–(13) (with or without covariation): people have their own sensory perception, so if two people see something, they each do so.

- (12) The children saw the photo.
- a. ✓**Distributive:** The children *each* saw the photo.
 - b. ✗**Nondistributive:** The children saw the photo jointly without each individually doing so.
- (13) The children saw a photo.
- a. ✓**Distributive without covariation:** There is one photo; each child saw it.
 - b. ✓**Distributive with covariation:** Each child saw a different photo.
 - c. ✗**Nondistributive:** There is one photo; the children saw it jointly but not individually.

Table 1.1 summarizes the way definite and indefinite objects interact with repeatable and non-repeatable actions to shape a predicate’s potential for a distributive understanding.

	Definite	Indefinite
Repeatable on obj. (<i>open</i>)	A&B opened the window. ✓ Dist : Each opened it. ✓ Nondist : Opened it together.	A&B opened a window. ✓ Dist : Each opened one. ✓ Nondist : Jointly opened one.
Not repeatable on obj. (<i>break</i>)	A&B broke the vase. ✗ Dist : Each broke it. ✓ Nondist : Jointly broke it.	A&B broke a vase. ✓ Dist : Each broke one. ✓ Nondist : Jointly broke one.

Table 1.1: (In)definiteness and (non)repeatability interact to constrain a predicate’s potential for a distributive understanding.

In light of the way definiteness constrains distributivity, it is often valuable to use indefinite objects in example sentences, to avoid limiting the understandings available to verbs describing actions that cannot be repeated on the same object: using an indefinite object allows *break a vase* to have a distributive understanding which would be unavailable with a definite object. But it is also good to use definite objects, to set aside covariation when it is not relevant, and as a reminder that distributivity and covariation are in principle distinct (which is why I have chosen *open the window* as a key example of a predicate that can be understood both distributively and nondistributively). When using definite objects, it is important to consider whether the action described by the verb can be repeated on the same object.

1.3.4 What’s possible versus what’s preferred

In experimental work on distributivity spanning multiple Indo-European languages (Brooks & Braine 1996, Frazier et al. 1999, Kaup et al. 2002, Dotlačil 2010, Pagliarini et al. 2012, Syrett & Musolino 2013, Dobrovie-Sorin et al. 2016, Maldonado et al. 2017), researchers have found that when a predicate can be understood both distributively and nondistributively (‘collectively’), then its nondistributive understanding is strongly preferred. For example (Dobrovie-Sorin et al. 2016), following a French sentence of the form *the children built a sand castle*, experimental participants choose to refer to the sand castle in the singular (so that all the children worked together to build a single sand castle — nondistributive), rather than in the plural (indicating multiple sand castles — distributive).

Most of these experiments are based on predicates with potentially covarying indefinite objects

(*build a sand castle*), not ones with non-covarying definite objects (*open the window*; (8)). It would be interesting to see if the distributive understanding of *open the window* is more available than that of *build a sand castle*. Then we could assess whether the observed dispreference for distributive understandings is really a dispreference for covarying indefinites.

Like most work in the semantics literature, this dissertation acknowledges that some understandings are preferred over others, but focuses on which ones are *possible* (see [Champollion to appear](#) for discussion). Ideally, we want to understand both issues, but they are conceptually distinct; preferences are gradient, while the question of whether an understanding is possible or not is arguably binary. Moreover, we can only rank one understanding as preferred over another when we already know that both are possible: we can only discover that the nondistributive understanding of *build a sand castle* is preferred when we already establish that it can be understood both distributively and nondistributively. That is why this dissertation focuses on possible understandings first, with preferences as an important future direction.

1.4 Outline of the dissertation

Chapter 2 asks what distributivity should be contrasted with. Generally, the antonym of ‘distributive’ is ‘collective’, but different authors define that term in different ways — some viewing it as the absence of distributivity, others in terms of inferences about collaboration and joint responsibility. Some authors posit a further distinction between ‘collective’ and ‘cumulative’ understandings, while others conflate these concepts. (The word ‘cumulative’ is used in two related but different ways in this literature — to characterize readings / understandings of sentences, and to characterize predicates that are closed under sum formation.) Based on an analysis of verbs with incremental objects ([Tenny 1987](#), [Krifka 1989](#), [Dowty 1991](#)) — objects whose parts correspond to parts of the event, such as *eat the pizza* — this chapter offers evidence for the view that collective and cumulative understandings should not be assigned distinct semantic representations, and that distributivity should simply be contrasted with nondistributivity.

Chapter 3 investigates where distributivity comes from, semantically or pragmatically. While acknowledging that many analyses from the literature capture the facts, this chapter presents a unified analysis where any predicate applied to a plural is individually true of each cell of a pragmatically determined cover (set of subparts) of the subject (Higginbotham 1981, Gillon 1987, Schwarzschild 1996), with different covers being entertained depending on what is known about the event described by the predicate. For *smile*, the only sensible cover places each member of the subject in their own cell (distributive). For *meet*, the only sensible cover places multiple individuals in the same cell (nondistributive). Given that people can open windows individually or jointly, hearers entertain both a cover placing each person in their own cell (distributive), and one placing both people in the same cell (nondistributive). The next step is to explain how the cover gets set for different predicates.

Chapter 4 investigates how the cover is set, by exploring which verb phrases are understood in which way(s). To ground wide-ranging quantitative claims, this chapter presents a dataset of ratings for the distributivity potential of over 2300 verb phrases (Glass & Jiang 2017). Verbs are separated into categories using the meaning-based system of Levin 1993, with objects for transitive verbs chosen from the Corpus of Contemporary American English (Davies 2008). This dataset provides evidence consistent with a series of theoretically motivated hypotheses:

- TRANSITIVE / INTRANSITIVE HYPOTHESIS: Predicates built from many intransitive verbs (*smile*) are only understood distributively, while those built from many transitive verbs (*open the window*) can be understood nondistributively as well as distributively (Link 1983, Glass 2017).
- BODY / MIND HYPOTHESIS: Predicates describing bodily or mental actions (*jump, meditate, swallow a pill, see a photo, like a book*) are understood distributively, given that individuals have their own bodies and minds and so can only carry out these events individually.
- MULTILATERAL HYPOTHESIS: Predicates describing inherently multilateral actions (*meet*) are understood nondistributively, given that individuals cannot carry out such actions alone.

- CAUSATIVE HYPOTHESIS: Predicates built from causative verbs (describing an event where the subject causes the object to change, such as *open the window*) can be understood nondistributively (as well as, perhaps distributively, depending on definiteness and repeatability), given that the nature of causation allows that multiple individuals' contributions may be jointly sufficient but individually insufficient to cause a result.
- INCREMENTAL HYPOTHESIS: Predicates with incremental objects (those whose parts correspond to parts of the event described by the predicate, such as *eat the pizza*) can be understood nondistributively (as well as, perhaps, distributively), given that multiple individuals might each carry out the event described by the verb on a different portion of an incremental object (eat a different part of a pizza), only jointly adding up to (eating) the whole thing.

By illuminating which understandings are entertained for which predicates, this chapter helps to make the cover analysis explanatory.

Chapter 5 turns from verbal predicates to adjectival ones. Among adjectives too, some are understood distributively (*the boxes are new* conveys that each box is new), some nondistributively (*the boxes are connected* conveys that they are mutually so), some in both ways (*the boxes are heavy* could convey that each box is heavy, or that they are individually light but jointly heavy). In addition to these three types of adjectives, there is also a fourth type: those that could plausibly have a nondistributive understanding, but actually strongly prefer to be understood distributively (Schwarzschild's 'stubbornly distributive' predicates; Schwarzschild 2011, Quine 1960). For example, *the boxes are tall* could conceivably mean that the boxes are tall as a stack while being individually short; but actually is taken to convey that each box is individually tall. Rather than viewing these 'stubbornly distributive' predicates as lexically idiosyncratic, this chapter follows Scontras & Goodman 2017 in explaining this preference pragmatically.

Along the lines of using pragmatic reasoning to explain the distributivity potential of adjectives, this chapter also takes on a more general open question: which adjectives are understood in which ways, and why? Particularly, a predicate can only be considered 'stubbornly distributive' if it has

an imaginable (but unavailable) nondistributive understanding. But it is an open question which predicates have such an understanding (like *heavy* and *tall*) and which do not (like *new*).

Applying the reasoning already used in the realm of verbal predicates, the distributivity potential of adjectives is argued to stem from world knowledge about the properties they describe. Based on the assumption that gradable adjectives such as *heavy* maps an individual to its ‘degree’ along a scale (for *heavy*, the scale of weight), this chapter uses the tools of measurement theory (Stevens 1946, Suppes & Zinnes 1962, Krantz et al. 1971, Krifka 1989, Schwarzschild 2002, Schwarzschild 2006, Sassoon 2007, Sassoon 2010, Lassiter 2011, Solt 2015, Lassiter 2017) — a way of expressing how the measurement $\mu(x \oplus y)$ relates to the measurements $\mu(x), \mu(y)$ of its constituent parts — to characterize the understandings available to different types of adjectival predicates. The proposal is that a gradable adjective *A* has a plausible nondistributive understanding if $\mu(a \oplus b)$ can exceed $\mu(a)$ and $\mu(b)$ individually along the scale associated with *A*. Then the contextual standard for what counts as *A* can be set so that $a \oplus b$ exceeds it while *a* and *b* individually fall short of it — a nondistributive understanding of *A*.

Among adjectives as well as verb phrases, the goal is to articulate the aspects of world knowledge that determine how different predicates are understood; for a verb phrase, the explanation lies in the nature of the event it describes, while for a gradable adjective, it lies in the structure of the scale associated with the adjective.

Chapter 6 summarizes and situates the dissertation within a larger context. Broadly, this dissertation tackles a well-studied topic from its lesser-studied angle of lexical semantics, pursuing a theory of distributivity which makes large-scale empirical predictions. Because inferences about distributivity are fundamentally inferences about how individuals can participate in eventualities (events and states / properties; Bach 1986), it seeks an explanation within the nature of these eventualities. The idea that a predicate’s distributivity potential ‘depends on world knowledge’ becomes predictive when combined with an explanation of what knowledge matters and why.

Chapter 2

‘Collective’ vs. ‘cumulative’

This chapter argues that semantically, distributive understandings should simply be contrasted with nondistributive ones. A proposed semantic ambiguity between ‘collective’ and ‘cumulative’ understandings is called into question based on evidence from predicates with incremental objects (such as *eat the pizza*; Tenny 1987, Krifka 1989, Dowty 1991), so that a concept from lexical semantics illuminates a longstanding debate in the study of distributivity.

2.1 Introduction

To identify a distributive understanding of a predicate, the diagnostic criterion is clear: a predicate is understood distributively if it is inferred to be individually true of each member of a plural subject, as in (1a). It is less obvious what criteria should be considered essential to a nondistributive understanding such as (1b), also termed a ‘collective’ understanding. As explained by Champollion to appear, collectivity could be defined negatively, as the absence of distributivity — a view adopted by Roberts 1987, Verkuyl 1994, Link 1998a, Winter 2000, Kratzer 2007, and ultimately defended here.¹ Alternatively, collectivity could be defined positively, as the presence of certain in-

¹Defining ‘collectivity’ as the absence of distributivity, Verkuyl 1994 offers the memorable term ‘kolkhoz collectivity’ — when a predicate is true of its subject as a whole, but not of each part, just as a Russian *kolkhoz* (collective farm) is owned by a group but not by any of its members.

ferences about collaboration, group action, and joint responsibility (Landman 1996, Landman 2000, Champollion 2010, Champollion 2017).

- (1) Alice and Bob opened the window.
- a. ✓**Distributive:** Alice and Bob *each* opened the window.
 - b. ✓**Nondistributive / Collective:** Alice and Bob opened the window jointly without each individually doing so.

If distributivity is simply contrasted with ‘collectivity’ in the sense of nondistributivity, then the space is split in two — distributive and not. But if distributivity is contrasted with a positively defined notion of ‘collectivity’, then there is room for a multi-way distinction — distributive, collective, and something else. Therefore, it is the authors who define collectivity positively (Landman and Champollion) who posit a further distinction within the space of nondistributivity, between collective and cumulative understandings of predicates.

On this three-way split, some predicates are understood distributively (2a), some are understood collectively (2b), and some are understood cumulatively (2c). Cumulative understandings such as (2c) are said to arise when a sentence involves multiple plurals (in (2c), a plural subject and a numeral plural object), in such a way that neither scopes over the other. Further, while the ‘collective’ (2b) is said to entail that the children coordinated and are jointly responsible, the ‘cumulative’ (2c) is said not to entail any such collaboration.

- (2)
- a. **Distributive:** The children smiled \rightsquigarrow they *each* smiled.
 - b. **Collective:** The children opened the window \rightsquigarrow opened it jointly / collaboratively.
 - c. **Cumulative:** The children ate two pizzas \rightsquigarrow each ate some pizza; two pizzas were eaten in all.

For authors who advocate this three-way distinction, it is reflected semantically. The distributive (2a) involves either a distributive operator (essentially a silent version of *each*, discussed further in

Chapter 3) or a meaning postulate stating that if multiple people *smile*, they each do (again, see Chapter 3). In the collective (2b), the subject is mapped from a regular plural into a special sort of individual known as a ‘group’, using the group-forming operator \uparrow (Link 1983), so that there is a single opening-the-window event whose agent is the group \uparrow (*the children*). The cumulative (2c) is analyzed so that there is a ‘plural’ event of eating with the plurality *the children* as its agent, and the plurality *two pizzas* as its theme (Krifka 1992).

When these different semantic representations are assumed, we also derive three different readings for a single sentence such as (3). (3a) is derived using a distributive operator (essentially a silent version of *each*; Chapter 3). (3b) is derived when the group \uparrow (*the children*) serves as the agent of a single inviting event, of which *six adults* is the theme, and is said to entail that the children coordinated their actions and are jointly responsible for the inviting. (3c) is derived when there is a ‘plural’ inviting event with the plurality *three children* as its agent and the plurality *six adults* as its theme. (3c) is supposed to entail that each child invited some adult(s) and each adult was invited by some child(ren); but unlike (3b), it does not entail any collaboration among the children.

- (3) Three children invited six adults. adapted Landman 2000: 130
- a. **Distributive:** Three children each invited six adults.
(*up to 18 adults total, depending on overlap*)
 - b. **Collective:** Three children worked together to invite six adults.
 - c. **Cumulative:** Three children engaged in inviting, and six adults were invited in all.

Other authors reject this three-way split, analyzing (3b) and (3c) as two different ways that a single semantic representation of (3) could be true, and therefore simply contrasting distributivity with nondistributivity (eschewing a positive definition of collectivity). This view is the one ultimately defended in this chapter. Towards that conclusion, §2.2 presents arguments for and against the purported collective / cumulative distinction, siding with those who reject this distinction.

Next, setting up the argument from incremental-object predicates, §2.3 introduces an assumption that is widely used to handle cumulative understandings: the idea that verbs and thematic roles are

inherently cumulative — closed under sum formation, like plurals. (As explained below, the word ‘cumulative’ is used in this literature in two related-but-distinct ways: for readings / understandings of sentences, and for predicates that are closed under sum formation.) §2.4 then shows that this common analysis of cumulative understandings actually encompasses far more data than generally acknowledged — not just sentences with plural objects, but also those with singular objects that are construed as incremental in the sense of [Tenny 1987](#), [Krifka 1989](#), [Dowty 1991](#). The result is that many predicates traditionally analyzed as collective must now be considered ambiguous with a cumulative reading, creating a problematic explosion of readings. This argument from incremental-object predicates thus serves as a reason not to distinguish collective and cumulative readings, but rather to treat them as two different ways that a single non-distributive understanding can be true.

2.2 Should ‘collective’ be separate from ‘cumulative’?

This section lays out the arguments for and against a collective / cumulative distinction. The debate involves two related issues: whether collectivity should be defined positively or negatively (§2.2.1); and whether collectivity should be distinguished from cumulativeness (§2.2.2).

2.2.1 For and against defining collectivity positively

For a positive definition of collectivity The main proponents of defining collectivity in positive terms are [Landman 2000](#) and [Champollion 2010](#). For Landman (although not for Champollion), this commitment is tied up in a broader goal of analyzing distributivity and plurality as reflexes of one another (discussed further in Chapter 3, where I review the distributivity literature).

As a brief sketch, Landman draws a parallel between predicates that are understood distributively, such as *smile*, and singular count nouns, such as *child*. The idea is that *child* applies only to individual children such as *Alice* (‘atoms’), not pluralities or groups thereof. To be predicated of a plural, *child* must be pluralized using the plural-forming operator \star from [Link 1983](#), which yields the closure of a set under sum formation. If the atomic individuals *Alice* and *Bob* are in the

denotation of the singular *child* (4), then the plurality *Alice and Bob* ($Alice \oplus Bob$, in the Link-style analysis of plurals; see Link 1983) is in the denotation of the pluralized *children* (5) — logically, $\star child$. Conversely, if *Alice and Bob* is in the denotation of pluralized $\star child$, then *Alice* and *Bob* are each in the denotation of singular *child*. If *Alice* is a child and *Bob* is a child, then *Alice* and *Bob* are children, and vice versa (6).

$$(4) \quad \llbracket child \rrbracket = \{Alice, Bob\}$$

$$(5) \quad \llbracket \star child \rrbracket = \{Alice, Bob, Alice \oplus Bob\}$$

$$(6) \quad child(Alice) \wedge child(Bob) \leftrightarrow \star child(Alice \oplus Bob)$$

Landman extends this picture to predicates like *smile*. For Landman, *smile* is like *child* in that — as a fact about its lexical entry — it applies only to ‘atomic’ individuals such as *Alice*, not pluralities or groups. To be predicated of a plural, *smile* must be pluralized using \star , just like *child*. In this way, if the singular *smile* is true of *Alice* and of *Bob*, then the plural $\star smile$ is true of the plurality *Alice and Bob*, and vice versa, guaranteeing the two-way entailment in (9). The plural operator \star simultaneously makes *smile* plural and distributive, achieving Landman’s goal of framing distributivity and plurality as ‘two sides of one and the same coin’ (Landman 1989a: 590–591).

$$(7) \quad \llbracket smile \rrbracket = \{Alice, Bob\}$$

$$(8) \quad \llbracket \star smile \rrbracket = \{Alice, Bob, Alice \oplus Bob\}$$

$$(9) \quad smile(Alice) \wedge smile(Bob) \leftrightarrow \star smile(Alice \oplus Bob)$$

If distributivity and plurality are intimately linked, then collective readings — since they are not distributive — must not involve plurality; even though they often superficially involve a morphologically plural subject and plural verb agreement, they must be basically singular. On this reasoning, Landman analyzes the collective understanding of (10) so that the un-pluralized predicate *open the window* applies not to *the children* as a plurality, but rather to *the children* as a ‘group’ — a special

sort of singular individual, similar to a ‘group noun’ such as *committee* (10b), derived via the group-forming operator \uparrow of Link 1983. The distributive understanding of (10) is derived when *open the window* is pluralized with \star and applied to *the children* as a plurality, so that *open the window* is individually true of each child (10a). Whereas *smile* takes only atomic individuals in its denotation, *open the window* is assumed to take both atomic individuals and groups, making (10) ambiguous between the plural, distributive (10a) and the singular, collective (10b).

- (10) The children opened the window.
- a. **Distributive:** \star *open the window*(*the children*)
 - b. **Collective:** *open the window*(\uparrow (*the children*))

On this system, distributive predication is equivalent to plural predication (involving \star , which simultaneously makes a predicate plural and distributive), while collective predication is equivalent to singular, group predication (involving the group-forming \uparrow).

For the collective \uparrow operator to be meaningful, Landman believes that collective predication must not become ‘a plural waste-paper basket’ (Landman 2000: 169), but instead should be identified positively by the presence of certain inferences — termed ‘thematic implications’ on the grounds that they arise when a thematic role such as *agent* is occupied by a group rather than a purely atomic individual such as *Alice*. Landman gives three examples of these thematic implications: collective responsibility, collective action, and collective body formation.

(11) (from Roberts 1987, who in turn credits Greg Carlson) is used to illustrate the thematic implication of **collective responsibility**, attributing the invasion not just to some rogue Marines, but to the Marine Corps as an organization, even the members who did not directly participate.

- (11) The Marines invaded Grenada. Roberts 1987: 147, who credits G. Carlson

(12) is said to imply **collective action**, conveying that the children coordinated their actions.

- (12) The children carried the piano upstairs. adapted Landman 2000: 166

As another example of collective action, [Champollion 2010](#) claims that *all the girls built the raft* ‘entails that the girls coordinated their actions and were jointly responsible for the result’ ([Champollion 2010](#): 223).

Finally, (13) is supposed to illustrate the thematic implication of **collective body formation**. If (13) describes a situation where the children have built a human pyramid, it can be used even if not every child touches the ceiling, but only the child at the top of the pyramid. Landman argues that (13) is parallel to a sentence with a singular subject: just as *Alice touches the ceiling* can be used when only Alice’s hand touches the ceiling, (13) can be used when only part of *the children* as a group (the highest-up child) touches the ceiling. For Landman, (13) shows that *the children* form a ‘collective body’.

(13) The children touch the ceiling. adapted [Landman 2000](#): 165

To sum up: when collectivity is defined positively, it is said to be associated with inferences about group action and responsibility, which are derived when a group (formed via \uparrow) fills the thematic role (e.g., ‘agent’) associated with the subject of that predicate.

Against a positive definition of collectivity Of course, this account is vulnerable to objections, particularly surrounding the thematic inferences said to arise when a group such as \uparrow (*Alice* \oplus *Bob*) fills a thematic role such as ‘agent’. As a technical point, [Magri 2012](#) objects to analyzing the nondistributive understanding of *the children opened the window* in such a way that *the children* forms a ‘group’; because then we would incorrectly predict *the children* to combine with predicates that exclusively apply to groups, as in the strange sentence *?the children have ten members*. And in general, [Verkuyl 1994](#) warns against using the label ‘collective’ ‘sloppily’ (p. 53), arguing that quantificational notions such as distributivity and nondistributivity must not be confused with elusive concepts of ‘togetherness, joint intention, and spatio-temporal proximity’ (p. 73).²

²Historically, even sentences such as *The children walked* — where each child is inferred to have walked (distributive) — were characterized as ‘collective’ in a situation where the children walked in a socio-spatially coordinated activity. For example, [Bartsch 1973](#) analyzes *Three men entered* as semantically ambiguous between a reading where they entered together (‘collective’) and one where they entered separately (‘distributive’), even though there is no doubt that if three

More specifically, these 'thematic inferences' are not well defined (Landman 2000: 169, Champollion 2010: 225). Landman even describes them as 'non-inductive', or non-logical (Landman 2000: 171). It is rather unusual for such non-logical inferences to be derived from the logical representation of a sentence. In fact, there is evidence that the inferences which Landman associates with collective / 'group' predication should be explained pragmatically instead.

For example, perhaps (11) attributes 'collective responsibility' to the Marine Corps as an organization not because its agent is the group \uparrow (*Marines*), but rather because we know that the Marines are a cohesive organization which carries out operations planned from the top (Roberts 1987: 147).

Turning to 'collective action', it is true that (12) conveys that the children undertook a 'joint action', but other predicates that would be analyzed as 'collective' lack this inference. (14) is presumably collective (at least, it is not distributive, because *wrote the Elements of Style* is not individually true of each person; nor is it cumulative in the sense of involving multiple plurals, since the object is singular). But despite being 'collective', (14) describes a situation in which Strunk and White did not collaborate, because E.B. White actually wrote a book expanding a leaflet written by his deceased English professor William Strunk. (One could describe this situation as collaboration or collective action, but then those terms become rather meaningless.)

(14) Strunk and White wrote *The Elements of Style*.

(14) shows that, even when the agent role is presumably filled by a group formed with \uparrow on Landman's assumptions, the 'thematic implication' of collective action may be absent. In the reverse direction, there are also examples in which a predicate is understood distributively (meaning that there is no collective / group predication), but we still draw inferences about collective action. (15) is distributive: if two people go running, they each do so. And yet, because the subject is *Maria and her husband* (who presumably often coordinate their activities), we defeasibly infer that they went running together, in a coordinated effort.

people enter a room, they each do so (distributive). I agree with Verkuyl that when interpersonal coordination is conflated with nondistributivity in this way, the issue is confused.

(15) This morning, Maria and her husband went running.

Instead of explaining elusive inferences about collaboration and responsibility in terms of a semantic notion of collective / 'group' predication via \uparrow , I argue that such inferences should be handled pragmatically.

2.2.2 For and against a collective / cumulative distinction

For a collective / cumulative distinction When collectivity is defined positively, collective understandings are contrasted with 'cumulative' ones (16b)–(16c). As previewed above, cumulative understandings are said to involve sentences with multiple plurals, for example in the object as well as the subject. Cumulative understandings are not distributive (in (16c), the predicate *invited six adults* is not individually true of each member of the subject); but neither are they collective on the positive definition thereof, in that (16c) need not involve collaboration or joint action among the children.

- (16) Three children invited six adults. adapted Landman 2000: 130 (= (3))
- a. **Distributive:** Three children each invited six adults.
(*up to 18 adults total, depending on overlap*)
 - b. **Collective:** Three children worked together to invite six adults.
 - c. **Cumulative:** Three children engaged in inviting, and six adults were invited in all.

The original example of a cumulative understanding, from Scha 1981, is (17c). This understanding is not distributive, in that *use 5k U.S. computers* is not true of each Dutch firm (it is not distributive); but nor is it collective on the positive definition of collectivity, in that it does not convey that the six hundred Dutch firms work together in any way (indeed, they may not even be aware of one another's computer usage). Instead, (17c) simply reports an aggregated U.S.-Netherlands trade statistic.

- (17) Six hundred Dutch firms use five thousand American computers. [Scha 1981](#): 132
- a. **Distributive:** 600 Dutch firms each use 5k U.S. computers (3 million computers total).
 - b. **Collective:** 600 Dutch firms jointly use 5k U.S. computers.
 - c. **Cumulative:** 600 Dutch firms use U.S. computers, 5k computers are used in all.

As further evidence for the purported collective / cumulative distinction, Landman and Champollion point to cases where one of these two understandings is true and felicitous while the other is false or unavailable. Champollion argues that (18) only has a collective reading, not a cumulative one, because it suggests that the Afghans as a group are collectively responsible for sending an emissary. In contrast, he says that (19) (from [Kroch 1974](#)) has only a cumulative reading, not a collective reading, because 'there is no sense in which the men have collective responsibility for being married to the [women] above and beyond their individual responsibilities' ([Champollion 2010](#): 55).

- (18) The Afghans sent an emissary to the Americans. [adapted Champollion 2010](#): 54
- a. **Distributive:** Each Afghan sent an emissary to the Americans.
 - b. **Collective (preferred):** The Afghans as a group sent an emissary to the Americans.
 - c. **Cumulative (not easily available):** Every Afghan engaged in emissary-sending, and every American received an emissary.
- (19) These men are married to those women. [adapted Kroch 1974](#)
- a. **Distributive (implausible):** Each man is married to the women.
 - b. **Collective (implausible):** The men as a group are married to the women.
 - c. **Cumulative (preferred):** Each man is married to some woman, and each woman is married to some man.

Moving from summary to critique, it is worth noting that the examples claimed to be three-ways ambiguous (16)–(19) actually just show that *if* collectivity is defined positively, in terms of

inferences about collaboration and joint responsibility (which is contentious), *then* we need a third category — cumulative — to account for the understandings that are neither distributive nor collective on this positive definition. But without such a positive definition of collectivity, such sentences would not need three semantically distinct ‘readings’; instead the ‘collective’ and ‘cumulative’ understandings would just be two different ways that a nondistributive understanding could be true (Roberts 1987, Verkuyl 1994, Link 1998a, Kratzer 2007).

Landman presents a more involved argument for the collective / cumulative distinction, comparing sentences (20)–(22) which differ along two dimensions: whether the numeral in the subject is greater than the numeral in the object or vice versa; and whether the subject is *women* or *chickens*. (I adjust Landman’s exact numbers for simplicity.)

First, let us investigate the relative magnitude of the numerals in the subject and object. Landman begins by claiming that (20) can have neither a cumulative reading nor a collective one.

- (20) Five women gave birth to three children. adapted Landman 2000
- a. **Distributive:** Each woman gave birth to three children (15 children total).
 - b. **Collective (strange):** Five women as a group gave birth to three children.
 - c. **Cumulative (inconsistent):** Five women gave birth to children, and three children were born in all.

If (20) did have a cumulative reading, it would mean that five women gave birth to children, and three children were born in all — but that is not possible, Landman says, because if five women gave birth to children, then at least five children would need to be born. Landman chooses the visceral example *give birth* because, barring medical complications, if someone *gives birth*, then at least one baby is born. The number of babies born must therefore equal or exceed the number of people giving birth. Since (20) states that only three children were born to five women, it cannot be understood cumulatively.

Nor can (20) be understood collectively, Landman says, because it is difficult to conceptualize a group of women as being jointly responsible for a certain number of births (on the assumption

that the collective reading — involving a thematic role filled by a group formed via \uparrow — would convey joint responsibility). Thus, (20) can only be distributive; it has no available collective nor cumulative reading.

In contrast, Landman says, (21) can have a cumulative reading, because it can describe a situation in which each of the three women gave birth to at least one child, and five children were born in all. The cumulative reading of (21) makes sense because the number of babies born exceeds the number of people giving birth.

- (21) Three women gave birth to five children. adapted Landman 2000
- a. **Distributive:** Each woman gave birth to five children (15 children total).
 - b. **Collective (strange):** Three women as a group gave birth to five children.
 - c. **Cumulative (available):** Three woman gave birth to children, and five children were born in all.

Just like (20), Landman says, (21) cannot have collective reading — again, because it is difficult to conceptualize a group of women as being jointly responsible for a certain number of births. Unlike (20), however, (21) *does* have a cumulative reading in addition to a distributive one.

Adding the contrast between *women* and *chickens*, Landman then argues that (22) can have a collective reading that (20) and (21) lack, because an industrial battery of chickens can be considered collectively responsible for its egg production, even if not every chicken in the group lays an egg (more generally, Landman assumes that collective readings do not entail that every member of the subject directly participated in the event, while cumulative readings do have this entailment). (22) cannot have a cumulative reading for the same reason that (20) cannot: because if fifty chickens engaged in egg-laying, then at least fifty eggs would need to be laid, not just thirty. But unlike both (20) and (21), (22) *does* have a collective reading, on the grounds that chickens can be construed as jointly responsible for their egg production (since all the chickens in an industrial battery are expected to produce eggs), while women are not generally considered jointly responsible for their

childbirths.³

- (22) Fifty chickens laid thirty eggs. adapted Landman 2000
- a. **Distributive:** Each chicken laid 30 eggs (1500 eggs total).
 - b. **Collective (available):** The chickens as a group laid 30 eggs.
 - c. **Cumulative (inconsistent):** Each chicken engaged in egg-laying; 30 eggs were laid in all.

In sum, Landman's main data points are that:

- i It is strange to say that *five women gave birth to three children*, while it is less strange to say that *three women gave birth to five children* (varying the relative magnitude of the numerals).
- ii It is strange to say that *five women gave birth to three children*, while it is less strange to say that *fifty chickens laid thirty eggs* (varying the subject as either *women* or *chickens*).

Landman takes these data as a 'serious problem' (Landman 2000: 174) for any attempt to collapse collectivity and cumulativity, and 'a strong argument here that cumulative readings are in fact not collective readings' (*ibid*).

To explain (i) (that it is strange to say that *five women gave birth to three children*, while it is better to say that *three women gave birth to five children*), Landman argues that the nondistributive understanding of (21) is a cumulative reading, which (20) lacks because five women cannot cumulatively give birth to only three children (given that each woman would have to give birth to a different child, which would result in more than three children). In contrast, if (21) were analyzed to have a collective reading, then (20) would be predicted to also have a collective reading, which it does not, because (20) cannot be nondistributive at all.

³Landman acknowledges that it is sometimes possible for the number of women to exceed the number of childbirths just as the number of chickens exceeds the number of eggs, as in 'hospital statistics' (Landman 2000: 174) such as 'our town's 10,000 women gave birth to 500 babies this year'. He analyzes these cases as collective, just like (22), so that 'this town's 10,000 women' would be construed as a 'group' occupying the thematic role of 'agent' of a single 'give birth' event (with '500 children' as the theme).

To explain (ii) (that it is strange to say that *five women gave birth to three children*, while it is better to say that *fifty chickens laid thirty eggs*), Landman says that neither (20) nor (22) can have a cumulative reading (because the number of offspring-producers exceeds the number of offspring), but that (22) can have a collective reading that (20) lacks because chickens but not women can be construed as collectively responsible for their offspring.

Against a collective / cumulative distinction However — again transitioning from summary to critique — there are ways of explaining Landman’s data in general terms, without positing an ambiguity between collective and cumulative ‘readings’. To explain (i) (that it is strange to say that *five women gave birth to three children*, while it is better to say that *three women gave birth to five children*), we might say that it is pragmatically odd to specify that five women gave birth to three children, since if only three children were born (each to only one woman), it is not clear how all five women participated in this event. It is less odd to specify that three women gave birth to five children, because each woman may participate in the event by giving birth to some of those children. This analysis does not actually call for any distinction between collectivity and cumulativity.

To explain (ii) (that it is strange to say that *five women gave birth to three children*, while it is better to say that *fifty chickens laid thirty eggs*), we might echo Landman’s idea that it is pragmatically sensible to tally the eggs laid by a certain number of chickens, given that all chickens in a battery are expected to produce eggs; while it is usually pragmatically odd to tally the babies born to a certain number of women, given that women are not expected to produce specific numbers of children. Again, this explanation does not actually require any semantic collective / cumulative distinction. (Looking forward, see §3.3.3 for an attempt to capture Landman’s data using the semantic analysis proposed in Chapter 3.)

In contrast to Landman and Champollion, other authors argue against a semantic ambiguity between collective and cumulative ‘readings’ (Roberts 1987 citing personal communication with Barbara Partee; Link 1998b, Link 1998a, Kratzer 2007, Dobrovie-Sorin et al. 2016). One part of this argument is to reject the positive definition of collectivity (following the concerns raised in §2.2.1). When that definition is rejected, the ‘collective’ and ‘cumulative’ understandings are analyzed not in

terms of a semantic ambiguity, but rather as two different ways that a nondistributive understanding could be true.

As empirical evidence for this viewpoint, [Kratzer 2007](#) offers ellipsis data: that *The two boys lifted the two boxes and the two girls did too* is true in a situation 'in which two boys jointly lifted each of the two boxes [collective], but the two girls each lifted a different one of the two boxes on her own [cumulative]' (p. 16). On the assumption that a true semantic or syntactic ambiguity cannot be resolved in two different ways in an antecedent and its ellipsis site ([Zwicky & Sadock 1975](#)), Kratzer concludes that 'we are right in lumping together collective and cumulative interpretations in a single reading' ([Kratzer 2007](#): 16).

Link offers a theoretical argument for the same conclusion. For him, the collective / cumulative debate raises 'a methodological point of a quite general nature in linguistics here: Where exactly does the line of demarcation run between proper readings and mere models realizing a reading?' ([Link 1991](#) and its English translation [Link 1998a](#): Chapter 2). I find his answer convincing:

'Distributive predication has universal quantificational force and is thus equipped with a precise logical interpretation. By contrast, the collective mode is mostly vague and indeterminate. Thus the empirical line is drawn between the distributive vs. the non-distributive (the rest)' [Link 1998b](#): 179–180 (page number from reprint in [Link 1998a](#): Chapter 7).

In other words, a distributive understanding is easily identifiable, requiring the predicate to be individually true of each member of the subject. There is no similarly clear criterion for distinguishing 'collective' or 'cumulative' understandings. Therefore, Link says, we should focus on modeling the clear distinction between distributivity and nondistributivity, not the elusive distinction between collectivity and cumulativity.

To sum up: I have now presented the literature's arguments for and against a semantic distinction between collective and cumulative understandings of predicates, coming down on the side of those who reject this distinction.

2.3 Cumulativity of verbs and thematic roles

The next step is to present this chapter’s strongest argument against a collective / cumulative distinction, based on evidence from predicates with incremental objects. To set up that argument, I first introduce a technical assumption often used to handle cumulative understandings of predicates: the assumption that verbs, and in neo-Davidsonian event semantics, thematic roles, are inherently cumulative in the sense of being closed under sum formation.⁴

Any predicate P is cumulative in this sense if it fulfills the definition in (23): if P is true of a and true of b , then it is true of their mereological sum ($a \oplus b$). A mass noun such as *wine* is cumulative in this sense, because if the liquid in cup a is wine, and the liquid in cup b is wine, then the liquid in both cups together is also wine (Quine 1960; Champollion & Krifka 2015 for an accessible introduction).

- (23) P is cumulative iff: Quine 1960

$$P(a) \wedge P(b) \rightarrow P(a \oplus b)$$

As a side note, this sense of cumulativity is the converse of distributivity: for P to be distributive means that if it is true of the sum $a \oplus b$, then it is true of a and true of b . Although these two definitions are converses of one another, it is not necessarily true that every cumulative predicate is also distributive, or vice versa. (An example is shown shortly.)

- (24) P is distributive iff:

$$P(a \oplus b) \rightarrow P(a) \wedge P(b)$$

Using this definition of cumulativity (23), some authors argue that verbs (and thematic roles such as ‘agent’) should be considered cumulative. This assumption would guarantee that if *smile* is

⁴For background: in the neo-Davidsonian event semantics of Castañeda 1967, Higginbotham 1985, and Parsons 1990 (inspired by Davidson 1967 and connected to distributivity by Schein 1986, Schein 1993), predicates are analyzed to relate individuals to the roles they play in an event; *Alice smiled* is analyzed to mean that there is a smiling event e with Alice as its agent: $\exists e[\textit{smile}(e) \wedge \textit{agent}(e, \textit{Alice})]$.

true of Alice and true of Bob, then *smile* is also true of Alice and Bob as a plurality ($Alice \oplus Bob$). Using event semantics, if there is a smiling event e_1 with Alice as its agent, and a smiling event e_2 with Bob as its agent, then there is also a larger event e_3 (the sum of $e_1 \oplus e_2$), also a smiling event, whose agent is the sum of the agents of e_1 and e_2 — in other words, whose agent is $Alice \oplus Bob$.

Representing the extension of *smile* as a set of events (Davidson 1967, Bach 1986, Parsons 1990), where each event is given (following Kratzer 2007) as a tuple listing its label and its thematic roles, then if e_1 and e_2 are in the extension of *smile*, their sum $e_1 \oplus e_2$ is also in this set.

$$(25) \quad \llbracket smile \rrbracket = \{ \langle e_1, agent = Alice \rangle, \\ \langle e_2, agent = Bob \rangle, \\ \langle e_1 \oplus e_2, agent = Alice \oplus Bob \rangle \}$$

In other words, (25) guarantees that if Alice smiled and Bob smiled, then Alice and Bob smiled.

Before proceeding, I offer some clarifying notes. Terminologically, the assumption reflected in (25) is called ‘summativity’ by Krifka 1989, ‘cumulativity’ by Krifka 1992, and ‘lexical cumulativity’ by Kratzer 2007 (who extends it to lexical items beyond verbs) and Champollion 2010 *et seq.* I call it ‘the assumption that verbs and thematic roles are cumulative’.

Whatever it is called, this assumption is widely adopted: for example, by Scha 1981, Lasnik 1989, Schein 1993, Landman 1996 / Landman 2000 (in a sense — see the footnote), Brisson 2003, Champollion 2010. However, it is directly opposed to claims by Carlson 1998 and Siloni 2012 that verbs are lexically singular (denoting only singular events, unless syntactically pluralized), and contrary to the assumptions of Landman sketched above (§2).⁵ Landman (and Carlson and Siloni) assumes that *smile* acts like a singular count noun such as *child*, and must be simultaneously plu-

⁵I cite Landman on both sides of this debate because different elements of his views align with each side. In analogizing distributivity to plurality, he wants verbs to act like singular count nouns, which would mean that they are not cumulative (they become simultaneously plural, distributive, and cumulative thanks to the \star operator). And yet in Landman 1996 and Landman 2000, he suggests that the basic, unmarked form of a verb such as *sing* is the plural form, $\star sing$ — suggesting that verbs *are* cumulative. But even though Landman takes $\star sing$ as the unmarked form, the only way for $\star sing$ to apply to a plural subject such as *Alice and Bob* is for a singular version of *sing* to apply to each atom in that plurality. (He still posits a singular / ‘atomic’ version of *sing* in addition to the unmarked plural form — see Landman 2000: Lecture Six for discussion.) That is how he reconciles his analysis of distributivity with the assumption that verbs and thematic roles are cumulative.

ralized and made distributive using Link’s pluralization operator, \star , in order to apply to a plurality such as *Alice and Bob*. In contrast, when we assume that verbs and thematic roles are cumulative, we assume that verbs never act like singular count nouns, but always act like plurals in being cumulative. They do not need to be pluralized using \star ; in Kratzer’s terms, they are already lexically ‘born’ plural.

Representationally, [Kratzer 2007](#) and [Champollion 2010](#) (and 2017) reflect the cumulativity assumption by prefacing all verbs and thematic roles with Link’s pluralizing operator \star , as in $\exists e[\star smile(e) \wedge \star agent(Alice)]$, as a reminder that *smile* and *agent* are taken to be closed under sum formation. But this convention may be confusing. Among authors who do not assume that verbs and thematic roles are cumulative, the \star operator may indicate distributivity as well as plurality (as in Landman’s system, previewed in §2 above). But for authors who use \star to reflect cumulativity of verbs and thematic roles, \star is *not* meant to convey distributivity. These authors assume that all verbs are cumulative (e.g., if Alice smiled and Bob smiled, then Alice and Bob smiled), but they do *not* assume that all verbs are distributive (e.g., it is *not* necessarily true that if Alice and Bob met, then Alice met and Bob met). When we assume that verbs and thematic roles are cumulative, then a verb like *meet* is cumulative in the sense of (23), but not necessarily distributive in the sense of (24) (as promised, providing an example where (23) and (24) come apart). It is important to remember when reading this literature that \star is used in different ways by different authors, sometimes indicating *both* distributivity *and* plurality (the Landman-style system) and sometimes indicating *only* plurality, *not* distributivity (in a Champollion-style system).

Note also that the word ‘cumulative’ is used in two slightly different ways in the literature and in this chapter. Above (§2.2), it was used to describe readings of sentences, such as *three children invited six adults* (where three children engaged in inviting, and six adults were invited in all). Here, it is used to describe a property of predicates, defined in (23). These senses of ‘cumulative’ are distinct, but they are related: as [Krifka 1992](#) shows, cumulative understandings of sentences can be perspicuously handled using the assumption that verbs and thematic roles are cumulative.

On this assumption, if Alice eats one pizza and Bob eats another pizza, it follows that Alice and

Bob eat two pizzas. Technically: if the extension of *eat* includes an eating event with Alice as its agent and one pizza as its theme, and another eating event with Bob as its agent and a second pizza as its theme, then it also includes an eating event with Alice and Bob as its agent, and two pizzas as its theme. This composite event $e1 \oplus e2$ (the third line of (26)) is the ‘sum’ of the two constituent events (Alice eating one pizza, Bob eating another; assuming that the agent of a sum event is the sum of the agents of each constituent event, and likewise for the theme; see Krifka 1992, Champollion 2010). If $e1$ and $e2$ are in the extension of *eat*, then their sum $e1 \oplus e2$ is there too, because *eat* is closed under sum formation (cumulative).

$$(26) \quad \llbracket eat \rrbracket = \{ \langle e1, agent = Alice, theme = pizza_1 \rangle, \\ \langle e2, agent = Bob, theme = pizza_2 \rangle, \\ \langle e1 \oplus e2, agent = Alice \oplus Bob, theme = pizza_1 \oplus pizza_2 \rangle \}$$

In other words, (26) captures the natural result that if Alice eats a pizza and Bob eats another pizza, then Alice and Bob eat two pizzas total.

This setup also raises a possibility for the reverse inference: that if Alice and Bob eat two pizzas, then perhaps they each eat some amount of pizza, adding up to two pizzas between them. Concretely, *Alice and Bob ate two pizzas* informs us that the third line of (26) is in the extension of *eat*, which is compatible with the extension of *eat* also including separate eating events by Alice and by Bob which together add up to *two pizzas* (perhaps Alice ate 0.5 pizzas and Bob ate 1.5; perhaps they each ate one, as represented in the first two lines of (26); or any other way of dividing two pizzas between two people). Such a situation verifies the ‘cumulative’ understanding of *Alice and Bob ate two pizzas*: Alice and Bob each did some pizza-eating, and two pizzas were eaten in all.

Therefore, the assumption that verbs and thematic roles are cumulative (closed under sum) naturally derives the ‘cumulative’ understanding of such sentences — a point in its favor (Krifka 1992). (In contrast, if one does not assume that verbs and thematic roles are cumulative, then such understandings can for example be derived by pluralizing both the subject and the object of the verb using two different \star operators; Beck & Sauerland 2000.)

To recap: I have introduced a common, motivated assumption — that verbs and thematic roles are cumulative — which is used to derive cumulative understandings. That assumption provides the background for the current chapter’s argument from incremental-object predicates.

2.4 Evidence from predicates with incremental objects

Traditionally, only sentences with multiple plurals are considered eligible for a cumulative understanding (e.g., *three children invited six adults*). When a sentence has a plural subject and a singular object (*the children opened the window, the children ate a pizza*), it is assumed that the only available understandings are distributive and collective, not cumulative. If one assumes a semantic distinction between collectivity and cumulativeness, the result is that sentences with multiple plurals can be three-ways ambiguous (distributive, collective, cumulative), while sentences with only one plural are only two-ways ambiguous (distributive, collective).

However, if one adopts the widespread assumption that verbs and thematic roles are cumulative, this traditional picture is actually not accurate. Instead (as already hinted by [Krifka 1992: §6](#) and [Dobrovie-Sorin et al. 2016: 90](#)), certain predicates with *singular* objects are also predicted to have a cumulative understanding: those with objects construed as incremental ([Tenny 1987](#), [Krifka 1989](#), [Dowty 1991](#)): for example, *eat the pizza*, where each part of the pizza corresponds to a part of the event of eating it and vice versa.

What are incremental objects? At this stage, it is worth clarifying what is meant by an incremental object, because different authors use this term differently.

Building on insights from [Verkuyl 1972](#), [Tenny 1987](#) observes that certain objects ‘measure out’ the event described by the predicate, so that the boundedness of an apple ‘delimits’ (gives an endpoint to) the event of eating it. Tenny’s notion of ‘measuring out’ is quite broad: it is meant to apply even to predicates that do not have an object whose parts correspond to the parts of the event described by the predicate ([Tenny 1994](#)) — mainly other types of ‘accomplishment’ predicates in the sense of [Vendler 1967](#) (events that are both durative and telic, in contrast to punctual, telic

achievements, and durative, atelic *states* and *activities*). For example (Tenny 1994), in the resultative predicate *scrub the sink clean*, it is not the sink, but the cleanness of it, that is incrementally affected (improved) over the course of the cleaning event.

In a similar spirit, Rothstein 2001, Rothstein 2004, Rothstein 2012 assumes that all accomplishments in the sense of Vendler 1967 (durative, telic events) are inherently incremental. She broadens the notion of incrementality in order to encompass even those accomplishments which do not involve gradually affecting subparts of the object (*repair the computer*, *sing the baby to sleep* — where neither the computer nor the baby is incrementally affected over the course of the event; rather, these events involve gradual advancement towards a result state). Therefore, both Tenny and Rothstein define incrementality broadly.

Krifka 1989 (and 1992) proposes the more restricted formal definition that I adopt here (echoed in prose by Dowty 1991, who introduces the term ‘incremental theme’). Krifka captures incrementality (‘graduality’) in terms of two symmetrical properties — ‘mapping to objects’ (27) and ‘mapping to events’ (28) — which together ensure a homomorphism between the parts of the object and the parts of the event. *Eat a pizza* has an incremental object because every part of the pizza-eating event corresponds to a part of the pizza (mapping to objects), and every part of the pizza corresponds to a part of the event of eating it (mapping to events).

(27) **Mapping to objects**

- a. $\forall R[MapObjects(R) \leftrightarrow \forall e, e', x[R(e, x) \wedge e' \sqsubseteq e \rightarrow \exists x'[x' \sqsubseteq x \wedge R(e', x')]]]$
- b. **Prose:** ‘A thematic role R has the mapping-to-objects property iff, given an event e in which an object x serves in the thematic role R , then for every subpart e' of the event e , there is a subpart x' of x which serves in the thematic role R of e' .’
- c. **Example** (adapted Krifka 1992: 39): ‘Every part of an eating-the-pizza event corresponds to a part of the pizza’.

(28) **Mapping to events**

- a. $\forall R[MapEvents(R) \leftrightarrow \forall e, x, x'[R(e, x) \wedge x' \sqsubseteq x \rightarrow \exists e'[e' \sqsubseteq e \wedge R(e', x')]]]$

- b. **Prose:** 'A thematic role R has the mapping-to-events property iff, given an event e in which an object x serves in the thematic role R , then for every subpart x' of x , there is a subpart e' of e such that x' serves in the thematic role R of e' .'
- c. **Example** (adapted Krifka 1992: 39): 'Every part of the pizza being eaten corresponds to a part of the eating event'.

This definition encompasses both physical events (*mow the lawn*) and mental ones (*read the book* — incremental because each portion of the book corresponds to a part of the event of reading it and vice versa; Krifka 1992: 44). But while all incremental-object predicates are accomplishments in the sense of Vendler 1967 (durative and telic, with their telicity derived from the boundedness of their objects as laid out by Krifka 1992), not all accomplishments are incremental. *Sing the baby to sleep* is an accomplishment without an incremental object, because the parts of the singing event do not map onto parts of the baby.

Adopting this definition, an incremental object is one whose parts correspond to the parts of the event described by the predicate: *the pizza* constitutes an incremental object in a sentence such as *Alice ate the pizza* because the progress of the event of eating the pizza mirrors the amount of pizza that is consumed.

Incremental-object predicates can be understood cumulatively Combining this incremental mapping between objects and events with the assumption that verbs and thematic roles are cumulative, the result is that when a sentence has a plural subject and a bounded incremental object (29), it is predicted to have a cumulative understanding in addition to the distributive and collective understandings already assumed.

(29) Alice and Bob ate the pizza.

- a. **Distributive (implausible given that the same pizza cannot be eaten twice):** They each ate the pizza.
- b. **Collective:** They ate the pizza jointly / collectively.

- c. **Cumulative:** Each did some pizza-eating; in total, the whole pizza was eaten.

The cumulative understanding (29c) is available because (29) asserts that the extension of *eat* includes an event of eating with *Alice and Bob* as its agent and *the pizza* as its theme (the third line of (30)). One way for this to be true is for the extension of *eat* to also include an event of eating with Alice as its agent and part of the pizza as its theme, and another event of eating with Bob as its agent and the rest of the pizza as its theme — adding up to an event of Alice and Bob eating the whole pizza between them:

$$(30) \quad \llbracket \textit{eat} \rrbracket = \{ \langle e_1, \textit{agent} = \textit{Alice}, \textit{theme} = \textit{half the pizza}_1 \rangle, \\ \langle e_2, \textit{agent} = \textit{Bob}, \textit{theme} = \textit{half the pizza}_2 \rangle, \\ \langle e_1 \oplus e_2, \textit{agent} = a \oplus b, \textit{theme} = \textit{half the pizza}_1 \oplus \textit{half the pizza}_2 \rangle \}$$

Thanks to the object-event mapping, the same reasoning used for predicates with numeral plural objects also extends to predicates with singular, incremental objects (compare (26) and (30)). It is not just sentences with multiple plurals that are eligible for a cumulative understanding (e.g., a plural subject and a plural object), but also sentences with plural subjects and *singular* objects that are construed as incremental.⁶

As a result, if one assumes a semantic ambiguity between collective and cumulative understandings, one must accept that this ambiguity is far more pervasive than generally imagined. Ultimately, I view this proliferation of ambiguity as an argument against the purported distinction between collective and cumulative readings.

⁶Other authors have also noted that incremental objects behave like numeral plurals in allowing a ‘cumulative reading’: namely Krifka 1992, Landman 2000, and Dobrovie-Sorin et al. 2016. Krifka 1992 uses the assumption that verbs (and thematic roles) are cumulative to handle incremental-object predicates (*eat a pizza*), and then extends the same analysis to numeral plurals such as *see seven zebras*. Dobrovie-Sorin et al. 2016 (p. 84 footnote 3; p. 90) point out that *the children built the sand castle* could be considered both ‘collective’ and ‘cumulative’ simultaneously (if the children work together to build a sand castle by each building a different portion of it) — briefly suggesting that this distinction is suspect, as I argue here. Landman 2000 (Lecture Six) observes that a sentence such as *The child ate a pizza* (adapted from his example, *a boy ate a bread*) can be represented ‘cumulatively’, as a sum of eating events of different portions of a pizza, adding up to a whole pizza in all. This reading is derived when an optional ‘mass partition’ operator is applied to a *pizza* — ‘a subtle shift of meaning of *eat*, focusing on the actual process of eating’ (p. 215). But Landman does not take this possibility as evidence against the proposed collective / cumulative distinction.

For example, if one assumes distinct semantic representations for distributive, collective, and cumulative understandings, (31) must now be considered three-ways ambiguous. The distributive (31a) would be derived from the presence of a distributive operator (silent *each*; discussed further in Chapter 3). The collective (31b) would be derived when the group \uparrow (*Alice* \oplus *Bob*) fills the thematic role of ‘agent’, creating ‘thematic implications’ of collective responsibility and collaboration. Finally, (31c) would be derived purely from the assumption that verbs and thematic roles are cumulative.

- (31) Alice and Bob painted the wall.
- a. **Distributive:** They each painted the wall.
 - b. **Collective:** They worked together to paint the wall.
 - c. **Cumulative:** They each did some painting, and the whole wall was painted in all.

But if Alice and Bob painted the wall collaboratively by each painting a different portion of it, then both (31b) and (31c) are true — so it is not clear whether the group-forming \uparrow operator should be present or not.

Along the same lines, one of the literature’s most often-repeated examples of a collective understanding — *the children built a raft* — must also be considered semantically ambiguous between a collective understanding (where the children worked together) and a cumulative one (where each child built a different part of the raft), given that *build a raft* can be construed as an incremental-object predicate.

Facing this proliferation of ambiguity (which does not really act like ambiguity anyway, at least with regard to Kratzer’s ellipsis test), along with the difficult task of distinguishing scarcely-different collective and cumulative understandings such as (31b) and (31c), there is a simple way out. We can reject the purported ambiguity, instead analyzing collective and cumulative understandings as two different ways that a nondistributive understanding of the sentence could be true. Rather than being derived when a group such as \uparrow (*Alice* \oplus *Bob*) serves in a particular thematic role, inferences about collaboration and joint responsibility could be explained pragmatically, based on our knowledge

about the cohesiveness of the subject and the nature of the event described by the predicate. That is what I propose to do here.

The goal of the following chapter (Chapter 3) is to present a semantic analysis of distributivity. Helping to delineate that task, the current chapter has argued that the semantic analysis should not model a three-way distinction between distributivity, collectivity, and cumulativity, but instead should just handle a two-way contrast between distributivity and nondistributivity.

2.5 Chapter summary

This chapter revisits a purported distinction between 'collective' and 'cumulative' understandings, arguing based on evidence from incremental objects that it is not needed. Instead, distributive understandings are just contrasted with nondistributive ones.

Chapter 3

Semantic representation

This chapter explores how distributive and nondistributive understandings should be represented semantically. While acknowledging that many analyses capture the data, I present a straightforward analysis in the spirit of [Higginbotham 1981](#), [Gillon 1987](#), and [Schwarzschild 1996](#): a predicate applied to a plural subject is individually true of each cell of a pragmatically determined cover — a set of subparts — of the subject. If each individual occupies its own cell of the cover, the predicate is understood distributively; if they all occupy the same cell, it is understood nondistributively. Inferences about distributivity are framed as inferences about which setting(s) of the cover to entertain, given what is known about the event described by the predicate.

3.1 Introduction

As illustrated above (Chapter 1), some predicates are understood distributively (1), some are understood nondistributively (2), and some can be understood in both ways (3).

- (1) The children smiled.
 - a. ✓**Distributive**: The children *each* smiled.
 - b. ✗**Nondistributive**: The children smiled jointly without each individually doing so.

- (2) The children met.
- a. **✗Distributive:** The children *each* met.
 - b. **✓Nondistributive:** The children met jointly without each individually doing so.
- (3) The children opened the window.
- a. **✓Distributive:** The children *each* opened the window.
 - b. **✓Nondistributive:** The children opened the window jointly without each individually doing so.

After sketching the data that needs to be captured (§3.2), this chapter presents an analysis which attributes all inferences about distributivity and nondistributivity to a single, fundamentally pragmatic source (§3.3). Applied to a plural, a predicate is required to be true of every cell of a pragmatically supplied cover of the subject. The setting of the cover is determined by how the members of the subject can participate in the event described by the predicate. This analysis explains very little on its own, but becomes explanatory when combined with a predictive theory of which predicates are understood in which ways (developed in Chapters 4 and 5).

Many other analyses (reviewed in §3.4) capture the same facts as the one I propose, so readers are invited to choose an alternative if they wish. I use the cover analysis only because I think it is the simplest, providing a transparent framework for investigating which predicates are understood in which ways (Chapters 4 and 5).

3.2 Data to capture

Much of the literature's discussion of distributivity centers on a handful of predicates. *Smile* exemplifies distributive predicates, *meet* or *gather* exemplify nondistributive ones. While I have used *open the window* to exemplify predicates that can be understood in both ways, a more common choice is *build a raft* (Link 1983) — understood so that only one raft is built on its nondistributive understanding, while multiple rafts (one per raft-building event) are built on its distributive

understanding (a distributive understanding ‘with covariation’; see §1.3.3). These exemplars are valuable; but it is equally important to apply a theory of distributivity to a broader range of data. So in reviewing each analysis, I investigate how it handles the following examples:

- *smile*
- *meet*
- *open the window*
- *build a raft*
- *lie* (in the sense of ‘mislead’)
- *see the photo*
- *smile* in an unusual context, applied to *lips* (Winter & Scha 2015)

By considering *open the window* in addition to *build a raft*, we observe how each analysis handles a definite, non-covarying object in addition to an indefinite, covarying one. Both predicates can be understood distributively and nondistributively, but *open the window* involves a single window which might be opened multiple times, while *build a raft* involves a different raft for each event of building one.

Like *open the window* and *build a raft*, the intransitive verb *lie* (in the sense of ‘mislead’) can also be understood in two ways (4) — distributively if each child lied, nondistributively if they lied in a jointly-issued statement.

- (4) The children lied.
- a. ✓**Distributive:** Each child lied.
 - b. ✓**Nondistributive:** The children lied jointly but not individually.

To exemplify predicates that can be understood in both ways, it is most common to use transitive verbs (*open the window*, *build a raft*). *Lie* tests how the theory handles both of these ways of understanding an intransitive verb.¹

Conversely, to exemplify predicates that are only understood distributively, it is most common to use intransitive verbs (*smile*). Like *smile*, *see the photo* is only understood distributively, in that if multiple people *see the photo*, they each do (5). Adding *see the photo* alongside *smile* shows how the theory handles this inference pattern for predicates built from transitive verbs as well as intransitive ones.

- (5) The children saw the photo.
- a. ✓**Distributive**: Each child saw the photo.
 - b. ✗**Nondistributive**: The children saw the photo jointly but not individually.

Another question is whether a predicate's distributivity potential is predicted to be rigid or flexible. However one explains that *smile* is distributive, one must also allow for unusual examples such as (6), which can arguably be understood nondistributively, given that lips can jointly create a smile in a way that humans cannot. I take (6) as further evidence that the distributivity of *smile* is not an arbitrary restriction on its lexical entry, but rather depends on the event it describes.

- (6) Alice's lips smiled (but her eyes didn't). adapted Winter & Scha 2015: 5
- a. (??) **Distributive**: Alice's lips *each* smiled.
 - b. ✓**Nondistributive**: Alice's lips smiled jointly.

By testing each theory of distributivity against these diverse predicates, elements are exposed which would remain hidden based only on *smile*, *meet*, and *build a raft*.

¹Similarly, de Vries 2015 discusses the two (distributive and nondistributive) understandings available to *win* (two people might each win different competitions, or might win a single competition jointly, for example in pairs figure skating). *Win* looks like an intransitive verb but may introduce confusion because it could also be analyzed to have a definite implicit object; Condoravdi & Gawron 1996.

3.3 A cover analysis

The proposed analysis is inspired by [Higginbotham 1981](#), [Gillon 1987](#), [Verkuyl & van der Does 1996](#), [Schwarzschild 1996](#), [Landman 1996](#), and in some sense [Moltmann 1997](#) and [de Vries 2015](#): that a predicate applied to a plural is individually true of each cell of a contextually supplied cover — set of subparts — of the subject.² I first review Schwarzschild’s version (§3.3.1), then introduce the version adopted here (§3.3.2).

3.3.1 Schwarzschild’s formulation

A cover ([Higginbotham 1981](#)) is defined as a set of subsets of a plural P (7).

- (7) C is a cover of P iff [Schwarzschild 1996](#): 64
- a. C is a set of subsets of P
 - b. Every member of P belongs to some set in C
 - c. \emptyset is not in C

For Schwarzschild, plurals are sets (while for those in the tradition of Link, plurals are sums such as *Alice* \oplus *Bob*; see [Lasersohn 2011](#), [de Vries 2015](#), [Champollion & Krifka 2015](#) for discussion of the differences). The set {a, b, c} has a number of different possible covers (8). Each subset of the initial set P is a cell. Each member of P could occupy its own cell (8a); the members of P could all occupy the same single cell (8b); two of the elements could be together in a single cell while the third is in its own cell (8c), and so on. The same element could even be represented in multiple cells, as in (8d). It is this possibility for repetition which distinguishes a cover from a more stringent notion known as a partition: a partition is a cover in which no element is represented in more than one cell, meaning that (8d) would not be permitted.

²[Moltmann 1997](#) argues that verb phrases are true of some contextually supplied part / whole structure of the subject, which is broadly similar to the cover analysis proposed here, although she formalizes her analysis with an unusual assumption that all verbs have ‘disjunctive’ (distributive and nondistributive) meanings. [de Vries 2015](#): Chapter 3 (p. 48) suggests that the distributive and nondistributive understandings of *win* can be attributed to different ways of identifying the relevant parts of the subject, similar to the cover analysis pursued here.

- (8) Covers of $\{a,b,c\}$
- a. $\{\{a\}, \{b\}, \{c\}\}$
 - b. $\{\{a,b,c\}\}$
 - c. $\{\{a, b\}, \{c\}\}$
 - d. $\{\{a,b\}, \{b,c\}\}$
 - e. ... (others) ...

On Schwarzschild's semantics, a predicate applied to a plural subject is separately true of each cell of a contextually supplied cover. (The cover is left as a free variable, to be saturated contextually like a pronoun; Schwarzschild specifically does not want it to be existentially quantified, because that would lead to very weak truth conditions). For every cell y of the cover of the plural subject x , the predicate α is required to be true of y , as given in (9). The Part operator provides universal quantification over all the cells in the cover.

- (9) $x \in \llbracket Part(Cov)(\alpha) \rrbracket$ iff
 $\forall y[(y \in Cov \wedge y \subseteq x) \rightarrow y \in \llbracket \alpha \rrbracket]$ Schwarzschild 1996: 71
 'A predicate α , given a contextually supplied cover, is true of a plurality x iff for every element y of the cover that is a subset of x , the basic predicate α is true of y '

For example, Schwarzschild observes that (3) can be understood to mean that each box is heavy (distributive), or that they are jointly heavy without each individually being so (nondistributive). (See Chapter 5 for more discussion of distributivity among adjectives.)

- (10) The boxes are heavy. adapted Schwarzschild 1996: 67
- ✓**Distributive:** Each box is heavy.
- ✓**Nondistributive:** The boxes are jointly heavy without each individually being so.

Both understandings can be derived from the semantics in (11): for every cell in the contextually supplied cover of the boxes, *heavy* is true of that cell. If the cover places each box in its own cell, as

in (8a), then we get a distributive understanding; if the cover places all the boxes in the same cell, as in (8b), we get a nondistributive understanding. Here, Schwarzschild suggests that the choice of cover depends on the discourse context: whether interlocutors care about the boxes individually or as a whole.

(11) The boxes are heavy.

$Part(Cov)(\llbracket heavy \rrbracket)(\llbracket boxes \rrbracket)$

‘For every element y of the contextually supplied cover which is a subset of *the boxes*, the basic predicate *heavy* is true of y ’

On this analysis, the two ways of understanding (11) do not correspond to a semantic ambiguity (Schwarzschild 1994, Schwarzschild 1996, Verkuyl & van der Does 1996, Moltmann 1997, Kratzer 2007, Nouwen 2015). Instead, there is only one semantics for (11), and the multiple ways of understanding it correspond to different pragmatic settings of the cover.

This analysis is designed to handle so-called ‘intermediate’ understandings of predicates: where the predicate is not true of each member of the subject, nor of the subject as a whole, but rather of some intermediate groupings. Describing a collection of, say, twenty shoes, (12) is not taken to convey that each shoe costs fifty dollars (distributive), nor that all the shoes together cost fifty dollars (nondistributive); but rather that each *pair* of shoes costs fifty dollars.

(12) The shoes cost fifty dollars.

Lasersohn 1998b: 88

Based on the knowledge that shoes are sold in pairs, the pragmatically supplied cover for (12) places each pair of shoes in its own cell.

While it seems like an advantage of the cover analysis that it can handle (12), some critics take it as a negative. Gillon 1987, Lasersohn 1989, Gillon 1990, and Lasersohn 1995 dispute whether sentences such as (13) should be predicted to have the intermediate understanding (13b) that the cover analysis allows.

- (13) *Context:* There are three Teaching Assistants (Alice, Bob, Caroline); Alice and Bob were each paid \$7,000 and Caroline was paid \$14,000.

Sentence: The TAs were paid \$14,000. adapted Lasersohn 1989: 131

- a. **Distributive:** The TAs were each paid \$14k (*false here*).

$$\text{Cov} = \{ \{a\}, \{b\}, \{c\} \}$$

- b. **Intermediate:** Two of the TAs were paid \$14k between them, the third was paid \$14k alone (*true here; but Lasersohn says this understanding is not available*).

$$\text{Cov} = \{ \{a, b\}, \{c\} \}$$

- c. **Fully nondistributive:** The TAs altogether were paid \$14k (*false here*).

$$\text{Cov} = \{ \{a, b, c\} \}$$

Lasersohn's position is that (13b) is not available, and that the cover analysis is wrong to predict it. Gillon replies that (13b) is available in a rich context — for example, where it is known that Caroline did twice as much work as Alice or Bob and so earned twice as much.

The disagreement between Lasersohn and Gillon points to a larger issue for this analysis: how speakers and hearers coordinate on the correct cover setting among many possible options. It is important to note that the cover analysis does not predict any imaginable cover to be available (contrary to what Lasersohn assumes); it has to be one that the speaker and hearer can coordinate on. To guide this coordination process, Schwarzschild proposes that speakers and hearers will avoid implausible, 'pathological' covers, such as $\{\{a, b\}, \{a, c\}\}$ for (13). Champollion 2016 suggests that out of context, the most available covers are the fully distributive one (placing each member of the subject in its own cell) and the fully nondistributive one (placing all the members in the same cell), because these options can be considered 'endpoints' (building on the Interpretive Economy Principle from Kennedy 2007, which is derived in terms of evolutionary game theory by Potts 2008;

Malamud 2006 also uses game-theoretic pragmatics to explain how interlocutors coordinate on the cover). Since the cover has to be one that the interlocutors can coordinate on, it is not surprising that some imaginable covers are unavailable. Nor does that fact constitute evidence against this analysis.

3.3.2 Analysis advocated here

Having reviewed Schwarzschild’s analysis, I present the revised version of it that I use here, beginning with the points of contrast between the original version and mine.

First, Schwarzschild does not use event semantics, and analyzes plurals as sets. To frame the analysis in the most widely used notation (although nothing hinges on these choices), I use event semantics, and I follow Link 1983 in taking plurals as sums rather than sets.

More substantively, Schwarzschild is motivated by handling ‘intermediate’ understandings such as the *shoes* example (12). He only uses the cover analysis where the predicate could plausibly be understood in multiple ways — distributively and nondistributively, like *build a raft*; or in some ‘intermediate’ way as in (12). He does not use it for predicates like *smile*, which he considers to be inherently distributive without any operators (based on the knowledge that people can only smile individually). In contrast, I see the cover analysis as a way to handle *all* inferences about distributivity and nondistributivity — not just ‘intermediate’ understandings (*the shoes cost \$50*) or cases where the predicate can be understood in multiple ways (*build a raft*), but also cases where it is only understood in one way (*smile, meet*).

Concretely, I analyze (14) to mean that each cell of the cover is the agent of a smiling event. Rather than assuming that *smile* is already inherently distributive, I derive its distributivity from pragmatic reasoning about the setting of the cover. Given that people can only smile individually, the only sensible cover is one that places each individual in their own cell (14a), yielding a distributive understanding. Diverging from Schwarzschild’s notation, $Cov(Alice \oplus Bob)$ is meant to return the set of cells of the contextually supplied cover of *Alice and Bob*.

(14) Alice and Bob smiled.

$$\forall x[x \in Cov(Alice \oplus Bob) \rightarrow \exists e[smile(e) \wedge agent(e, x)]]$$

- a. **✓Distributive:** they each smiled.
 $Cov = \{ \{a\}, \{b\} \}$
- b. **✗Nondistributive:** they smiled jointly but not individually.
 $Cov = \{ \{a, b\} \}$

Furthermore, I also require that the chosen cover must be the ‘tightest-fitting’ one — one where no cover with more fine-grained cells would be accurate. Without this stipulation, (14a) and (14b) would both be equally good covers in a situation where Alice and Bob each smiled — (14a) because each of them smiled, and (14b) because, when we assume that verbs and thematic roles are cumulative, then if there is a smiling event by Alice and a smiling event by Bob, there is also a larger smiling event by $Alice \oplus Bob$ (see §2.3). By requiring the tightest-fitting cover, (14a) is chosen over (14b) when Alice and Bob each smiled.

In the same way, (15) is assigned a cover placing Alice and Bob each in their own cell, on the grounds that people have their own sensory perception and so cannot see something jointly without also doing so individually.

(15) Alice and Bob saw the photo.

$$\forall x[x \in Cov(Alice \oplus Bob) \rightarrow \exists e[see(e) \wedge experiencer(e, x) \wedge theme(e, \iota y[photo(y)])]]$$

- a. **✓Distributive:** They each saw the photo.
 $Cov = \{ \{a\}, \{b\} \}$
- b. **✗Nondistributive:** They saw the photo jointly but not individually.
 $Cov = \{ \{a, b\} \}$

Next, (16) gets a cover placing Alice and Bob in the same cell, given that individual people cannot meet unilaterally.

(16) Alice and Bob met.

$$\forall x[x \in Cov(Alice \oplus Bob) \rightarrow \exists e[meet(e) \wedge agent(e, x)]]$$

- a. **✗Distributive:** They each met
Cov = { {a}, {b} }
- b. **✓Nondistributive:** They met jointly but not individually
Cov = { {a, b} }

Turning to predicates that can be understood both distributively and nondistributively, there are two covers available to (17) — a distributive one placing Alice and Bob each in their own cell (17a), and a nondistributive one placing them both in the same cell (17b) — based on the world knowledge that people can open windows individually or jointly.³

(17) Alice and Bob opened the window.

$$\forall x[x \in Cov(Alice \oplus Bob) \rightarrow \exists e[open(e) \wedge agent(e, x) \wedge theme(e, iy[window(y)])]]$$

- a. **✓Distributive:** They each opened it.
Cov = { {a}, {b} }
- b. **✓Nondistributive:** They opened it jointly but not individually.
Cov = { {a, b} }

The same goes for (18): two different covers are entertained, distributive and nondistributive, given that people can lie individually or in jointly issued statements.

(18) Alice and Bob lied.

$$\forall x[x \in Cov(Alice \oplus Bob) \rightarrow \exists e[lie(e) \wedge agent(e, x)]]$$

- a. **✓Distributive:** They each lied.
Cov = { {a}, {b} }
- b. **✓Nondistributive:** They lied jointly but not individually.

³If Alice and Bob opened the window jointly (nondistributive), then (17b) is the ‘tightest-fitting’ cover setting, because *open the window* is true of Alice and Bob together, but not individually true of each of them. If they each opened the window, then the ‘tightest-fitting’ cover is (17a).

$$\text{Cov} = \{ \{a, b\} \}$$

For (19), the distributive cover captures a situation in which they each build a different raft (distributive with covariation), while the nondistributive cover characterizes a situation in which they jointly build a single raft. Again, both covers are available thanks to our world knowledge that rafts can be built by individuals or by larger parties.

(19) Alice and Bob built a raft.

$$\forall x[x \in \text{Cov}(\text{Alice} \oplus \text{Bob}) \rightarrow \exists e \exists y[\text{build}(e) \wedge \text{agent}(e, x) \wedge \text{raft}(y) \wedge \text{theme}(e, y)]]$$

a. **✓Distributive (with covariation):** They each built a (different) raft.

$$\text{Cov} = \{ \{a\}, \{b\} \}$$

b. **✓Nondistributive:** They built a single raft jointly but not individually.

$$\text{Cov} = \{ \{a, b\} \}$$

Build a raft only has a distributive understanding with covariation (each person builds a different raft), given that the same raft cannot generally be built more than once. But the same semantics is also suitable for predicates that can be ‘distributive without covariation’. (20) is analyzed in the same way as (19), but the non-covarying scenario is captured if the existential quantifier in (21a) picks out the same photo for both Alice and Bob.⁴ That possibility does not make sense for (19), but is available for (20) given that the same photo can be seen multiple times.

(20) Alice and Bob saw a photo.

$$\forall x[x \in \text{Cov}(\text{Alice} \oplus \text{Bob}) \rightarrow \exists e \exists y[\text{see}(e) \wedge \text{experiencer}(e, x) \wedge \text{photo}(y) \wedge \text{theme}(e, y)]]$$

a. **✓Distributive (with or without covariation):** They each saw a (possibly different) photo.

$$\text{Cov} = \{ \{a\}, \{b\} \}$$

⁴Of course, there are other, perhaps better theories of indefinites that do not treat them as existential quantifiers — see McNally 1997, Reinhart 1997; and de Vries 2015 for a connection between such analyses and distributivity. But the simple existential quantifier analysis serves for current purposes.

- b. **✗Nondistributive:** They saw a single photo jointly but not individually.

$$\text{Cov} = \{ \{a, b\} \}$$

Finally, the *lips smiled* example (6) is understood nondistributively (where both lips occupy the same cell of the cover) based on the knowledge that lips can jointly create a smile.

- (21) Alice’s lips smiled.

$$\forall x[x \in \text{Cov}(\textit{lips}) \rightarrow \exists e[\textit{smile}(e) \wedge \textit{agent}(e, x)]]$$

- a. **(??) Distributive:** Each lip smiled.

$$\text{Cov} = \{ \{\textit{lip1}\}, \{\textit{lip2}\} \}$$

- b. **✓Nondistributive:** The lips smiled jointly but not individually.

$$\text{Cov} = \{ \{\textit{lip1}, \textit{lip2}\} \}$$

In other words, the cover analysis is grounded in the idea — which most researchers would agree with, in some form — that distributivity ‘depends on world knowledge’. Inferences about distributivity are inferences about which cover settings to entertain, given what is known about the event described by the predicate.

3.3.3 Capturing the ‘collective’ / ‘cumulative’ data on the proposed analysis

For completeness, I also sketch how this analysis — which does not semantically distinguish between ‘collective’ and ‘cumulative’ understandings — handles the data which Landman 2000 takes to motivate such a distinction (§2.2.2):

- i It is strange to say that *three women gave birth to five children*, while it is less strange to say that *five women gave birth to three children* (varying the relative magnitude of the numerals).
- ii It is strange to say that *five women gave birth to three children*, while it is less strange to say that *fifty chickens laid thirty eggs* (varying the subject as either *women* or *chickens*).

On the proposed analysis, (22) can get a cover placing each woman in her own cell (distributive: each woman gives birth to five children); or a cover placing all three women in the same cell (nondistributive: all three women jointly give birth to five children). (22) could also get ‘intermediate’ covers — for example, grouping {woman1, woman2} in the same cell and {woman3} in a different cell — but those would require a great deal of supporting context which is not available here.

(22) Three women gave birth to five children.

$$\forall x[x \in Cov(3 \text{ women}) \rightarrow \exists e[\text{birth}(e) \wedge \text{agent}(e, x) \wedge \text{theme}(e, 5 \text{ children})]]$$

a. **✓Distributive:** Each of 3 woman gave birth to 5 children (15 children total).

$$Cov = \{ \{ \text{woman1} \}, \{ \text{woman2} \}, \{ \text{woman3} \} \}$$

b. **✓Nondistributive:** All 3 women jointly gave birth to 5 children (5 children total).

$$Cov = \{ \{ \text{woman1}, \text{woman2}, \text{woman3} \} \}$$

The distributive cover (22a) makes sense: it is entirely possible for each of three women to give birth to five children — adding up to fifteen children total. The fully nondistributive cover (22b) also makes sense: three women were the agent of an event in which five children were born. Assuming that verbs and thematic roles are cumulative (§2.3), such an event could comprise component events in which woman1 gives birth to two children; woman2 gives birth to two children; and woman3 gives birth to one child — adding up to a ‘sum’ event of all three women giving birth to five children between them, which is the situation verifying the ‘cumulative’ understanding of (22). Both the distributive and nondistributive (‘cumulative’) understandings of (22) are derived, capturing Landman’s observation that (22) can be understood both distributively and nondistributively.

In general, when a sentence contains multiple plurals as in (22), one way for that sentence to be true on its nondistributive understanding is for each member of the plurality in the subject to carry out the event described by the verb on some part(s) or member(s) of the plurality in the object, between them adding up to the full object — as in (22b), where each woman gives birth to one or more of the children, adding up to five children in all. Such situations (represented by placing all

members of the subject in the same cell of the cover) verify the ‘cumulative’ understanding of the sentence, but without positing a semantic distinction between ‘cumulative’ and ‘collective’ representations. Non-cumulative ‘collective’ understandings (such as *Alice and Bob hired an employee*) are also represented by placing both members of the subject in the same cell of the cover, so ‘collective’ and ‘cumulative’ are not semantically distinct, even though they correspond to different types of situations.

Returning to Landman’s data, the next step is to explain the contrast between (22) (which can be understood both distributively and nondistributively) and (23) (only understood distributively). Like (22), (23) could hypothetically get a cover placing each woman in her own cell (distributive: each woman gives birth to three children); or a cover placing all five women in the same cell (nondistributive: all five women jointly give birth to three children); as well as various intermediate covers which I ignore. The distributive cover (23a) makes sense: five women could each give birth to three children. In contrast, the nondistributive cover (23b) is puzzling: it states that five women served as agents in an event in which three children were born, raising questions about how each woman participated in this event and why the speaker decided to mention all five of them when some of them didn’t have any children. Landman’s observed contrast between (22)–(23) is therefore explained without any semantic distinction between ‘collective’ and ‘cumulative’ readings.

(23) Five women gave birth to three children.

$$\forall x[x \in Cov(5 \text{ women}) \rightarrow \exists e\exists y[birth(e) \wedge agent(e, x) \wedge theme(e, 3 \text{ children})]]$$

a. **✓Distributive:** Each of 5 women gave birth to 3 children (15 children total).

$$Cov = \{ \{woman1\}, \{woman2\}, \{woman3\}, \{woman4\}, \{woman5\} \}$$

b. **✗Nondistributive:** All 5 women jointly gave birth to 3 children (3 children total).

$$Cov = \{ \{woman1, woman2, woman3, woman4, woman5\} \}$$

Finally, (24) can get a cover placing each chicken in its own cell (distributive: each chicken lays thirty eggs); or a cover placing all thirty chickens in the same cell (nondistributive: all fifty chickens jointly lay thirty eggs). The distributive cover again makes sense: each chicken lays

thirty eggs. The nondistributive cover states that fifty chickens were involved in an event in which thirty eggs were laid. As in (23), it is not clear how all fifty of these chickens participated in this event; presumably some of them did not lay any eggs. But this time we can more easily imagine why the speaker decided to mention all fifty of them (even those who did not lay any eggs): to aggregate the egg output of some relevant quantity of chickens. It makes sense in an industrial context to include non-egg-laying chickens in a tally of egg production, because all such chickens are expected to produce eggs; whereas it is pragmatically odd to include child-free women in a tally of childbirths, because such women are not generally expected to have children. Thus, Landman’s observed contrast between (23)–(24) is also explained pragmatically (along the lines that Landman himself suggests), without to any semantic distinction between ‘collective’ and ‘cumulative’.

(24) Fifty chickens laid thirty eggs.

$$\forall x[x \in Cov(50 \text{ chickens}) \rightarrow \exists e[lay(e) \wedge agent(e, x) \wedge theme(e, 30 \text{ eggs})]]$$

a. **✓Distributive:** Each of 50 chickens laid 30 eggs (1500 eggs total).

$$Cov = \{ \{chicken1\}, \{chicken2\}, \{chicken3\}, \{chicken4\} \dots \}$$

(each chicken in its own cell)

b. **✓Nondistributive:** All 50 chickens jointly laid 30 eggs (30 eggs total).

$$Cov = \{ \{chicken1, chicken2, chicken3, chicken4 \dots \} \}$$

(all chickens in the same cell)

In sum: the cover analysis does not distinguish semantically between collective and cumulative ‘readings’, but is argued to still capture the data taken to motivate such a distinction.

3.3.4 Discussion

This section has put forth a semantic analysis which simply requires a predicate to be individually true of each cell of a cover (set of subparts) of its plural subject.

Like the others reviewed below, this analysis does not on its own answer the question of which predicates are understood in which ways. It frames that question as a question about which cover(s)

to entertain, given what is known about the events described by various predicates; but it must ultimately be combined with a theory that answers that question. Of the alternatives reviewed in the following section, many capture the same data, so the proposed analysis is chosen only because I see it as the most straightforward one.

3.4 Alternative analyses from the literature

This section reviews some alternative analyses from the literature, organized by the number of distinct sources of distributivity assumed by each one. Setting aside a minority view which explains distributivity within the subject of the sentence (Bennett 1974, revived in Ouwayda 2014, Ouwayda 2017 based on Arabic data for syntactic reasons; see Dowty 1987, Lasersohn 1995, Champollion to appear for critique), I focus on the mainstream type of analysis: those which locate inferences about distributivity within the predicate.

3.4.1 One source: an operator

As previewed in Chapter 2 (§2.2.1), one way of handling distributivity is to analogize it to plurality, as in the work of Landman 1989a, Landman 2000, and those inspired by it. On this view, all distributive understandings are attributed to the pluralizing \star operator of Link 1983. To explain why *smile* is distributive, Landman argues that *smile* only applies to atomic individuals such as *Alice*, not groups or pluralities (25) (contrary to the assumption that verbs and thematic roles are cumulative; §2.3). To apply to a plurality, *smile* must be pluralized using the \star operator, which yields the closure of a set under sum formation (26). If un-pluralized *smile* is true of Alice and of Bob, then pluralized \star *smile* is true of the plurality $Alice \oplus Bob$, and vice versa (27), guaranteeing that *smile* is distributive. (For simplicity, the representations in (25)–(27) do not use event semantics.)

$$(25) \quad \llbracket smile \rrbracket = \{Alice, Bob\}$$

$$(26) \quad \llbracket \star smile \rrbracket = \{Alice, Bob, Alice \oplus Bob\}$$

$$(27) \quad \textit{smile}(\textit{Alice}) \wedge \textit{smile}(\textit{Bob}) \leftrightarrow \star\textit{smile}(\textit{Alice} \oplus \textit{Bob})$$

When *smile* is applied to *lips* rather than people, it would presumably have to apply to groups such as $\uparrow(\textit{lips})$, even though it otherwise only applies to atomic individuals.

See the photo would be handled like *smile* (the normal, predicate-of-people-not-lips version of it), guaranteed to apply only to atomic individuals.

In contrast, *meet* cannot apply to atomic individuals such as *Alice*, nor to pluralities such as *Alice and Bob*. *Meet* only applies to groups — special individuals made up of other individuals, including group nouns (*committee*), and groups derived by applying the group-forming operator \uparrow to a regular plurality, as in $\uparrow(\textit{Alice} \oplus \textit{Bob})$ (see Chapter 2). As a result, *meet* is nondistributive (‘collective’, in Landman’s terms). Moreover (Chapter 2), when a group such as $\uparrow(\textit{Alice} \oplus \textit{Bob})$ serves in the thematic role of ‘agent’, Landman says that we derive non-logical inferences about joint responsibility and collaboration.

$$(28) \quad \textit{meet}(\uparrow(\textit{Alice} \oplus \textit{Bob}))$$

For Landman, predicates like *open the window*, *build a raft*, and *lie* can apply to both atomic individuals and groups, so that they are ambiguous between a ‘plural’ distributive reading (derived using \star) and a ‘singular’ collective reading (using \uparrow), as in (29).

- (29) Alice and Bob opened the window.
- a. **Distributive:** $\star\textit{open the window}(\textit{Alice} \oplus \textit{Bob})$
 - b. **Collective:** $\textit{open the window}(\uparrow(\textit{Alice} \oplus \textit{Bob}))$

This analysis attributes all inferences about distributivity to a single source, parsimoniously connecting it to the related concept of plurality. Distributive and collective readings (and they are readings, for Landman) are analyzed in terms of a semantic ambiguity; collective readings evoke inferences about joint responsibility (a view critiqued in Chapter 2).

On such an analysis, the question of which predicates are understood in which ways is framed

as a question about which predicates take in their denotation atoms, groups, or both, which in turn presumably depends on what is known about the events described by those predicates.

3.4.2 One source: meaning postulates

A different approach attributes all distributive inferences to meaning postulates: restrictions on the models that we entertain in a model-theoretic framework, meant to represent world knowledge about the event described by the predicate (Champollion 2010: 159).

For Scha 1981, sentences such as (30) and (31) have the same semantics, but given what we know about *smiling* (formalized using a meaning postulate), we know that *smile* ‘trickles down’ to all the members of its subject, whereas *meet* does not. Whenever a multi-part entity G — a plurality or a group — smiles, generally every member of G also smiles. (Note that Scha implicitly assumes that verbs and thematic roles are cumulative [§2.3], in that *smile* applies not just to atomic individuals such as *Alice*, but also to sums such as $Alice \oplus Bob$. As above, (30) and (31) are not framed in event semantics, just for simplicity).

(30) Alice and Bob smiled.

$smile(Alice \oplus Bob)$

Meaning postulate: $smile(G) \rightsquigarrow \forall x[x \in G \rightarrow smile(x)]$

(31) Alice and Bob met.

$meet(Alice \oplus Bob)$

Scha does not handle unusual cases such as the *lips smiled* example, but presumably one would need a separate meaning postulate to handle *smile* when it is applied to *lips* versus people, to capture the knowledge that lips and people participate differently in *smile* events. Like *smile*, *see the photo* would also be associated with a meaning postulate guaranteeing that it is understood distributively.

One problem with this analysis is that it cannot handle predicates that are understood both distributively and nondistributively, such as *open the window*, *build a raft*, and *lie*. Meaning postulates

cannot be optional (Roberts 1987, de Vries 2015, de Vries 2017), so a meaning postulate cannot derive a distributive understanding of a predicate that can also be understood nondistributively.

Nor can the analysis handle ‘covariation’ (Dotlačil 2010) — cases where an operator in the verb phrase (an indefinite, a numeral, and so on) is applied separately to each member of the subject (Roberts 1987, Lasersohn 1995, Winter 1997, Winter 2000, Champollion 2010). For example, using only a meaning postulate attached to the verb *build*, it is not obvious how we could derive the understanding of (32) in which Alice and Bob each built a separate raft. We could propose a meaning postulate ensuring that whenever multiple individuals build something, they each do (which actually seems wrong, because in general, multiple individuals can jointly build something without each separately building the whole thing). But no meaning postulate can get *a raft* to covary with the members of the subject.

(32) Alice and Bob built a raft.

$$\exists y[\textit{raft}(y) \wedge \textit{build}(\textit{Alice} \oplus \textit{Bob}, y)]$$

Strange meaning postulate: $\textit{build}(G, y) \rightsquigarrow \forall x[x \in G \rightarrow \textit{build}(x, y)]$

To get (32) to mean that Alice and Bob each built their own raft (covariation), we would want *a raft* to take ‘narrow scope’ in some sense — but with respect to what? Without any other quantifier in the sentence, there is nothing for *a raft* to scope under. This problem is why many authors posit some sort of quantifier in their analysis of distributivity (see de Vries 2017 for discussion).

Because of these limitations, no current authors analyze all distributivity inferences in terms of meaning postulates, as Scha initially suggested. But the idea of meaning postulates still lives on, in approaches to distributivity which posit two distinct sources for it.

3.4.3 Two sources: meaning postulates and an operator

The most common approach to distributivity (Dowty 1987, Roberts 1987, Hoeksema 1988, Lasersohn 1990b, Lasersohn 1995, Link 1998a: Chapter 2, Winter 1997, Winter 2000, Winter 2002, de Vries 2015, de Vries 2017, Champollion 2010, Champollion 2017) is two-pronged. Predicates

like *smile* are handled using meaning postulates, intended to reflect that these predicates are distributive purely because of what we know about the events they describe. Predicates like *build a raft* are analyzed using an optional operator — sometimes known as the *D* operator (Link 1991, originally written in 1984, and its English translation Link 1998a: Chapter 2; Roberts 1987) and sometimes subsumed under the pluralizing \star operator of Link 1983 and Landman 1989a — which essentially acts like a silent version of *each*. *D* makes sure that the predicate is separately applied to each member of the subject.

This analysis can be seen as a synthesis of the two others we have seen: the meaning postulate approach of Scha for *smile*-type predicates, and the operator-based approach of Landman for *build a raft*-type predicates. The meaning postulate approach works for predicates like *smile*, which are (nearly) always understood distributively, but struggles to handle predicates that can be understood in both ways, particularly those where an operator (an indefinite, a numeral) in the verb phrase covaries with the members of the subject (*build a raft*, *make ten thousand dollars*). Accordingly, the meaning postulate approach is preserved where it is effective (for *smile*), while an optional operator is used to derive the distributive understanding of predicates that are optionally distributive (*build a raft*).⁵

⁵de Vries 2015 offers a further argument for the two-pronged approach based on group nouns such as *committee*. For Champollion 2010, Champollion 2016, Champollion 2017 (thanks to Lucas Champollion p.c. for discussion), the two-pronged approach is needed to explain why (ia) (adapted from Gillon 1987) is a true, felicitous description of the scenario in (i), whereas (ib) is judged false unless a favorable context already groups the musicians into pairs. (In reality, Richard Rodgers and Oscar Hammerstein co-wrote *many* musicals, as did Richard Rodgers and Lorenz Hart; but pretend for the moment that each duo only co-wrote one musical.)

- (i) *Scenario*: Rodgers and Hammerstein co-wrote a musical; Rodgers and Hart co-wrote a musical.
- a. (*Judged true, needs no additional context*.) Rodgers, Hammerstein, and Hart wrote musicals.
 - b. (*Judged false in this scenario without contextual support*.) Rodgers, Hammerstein, and Hart wrote a musical.

For Champollion, (ia) is true because verbs and thematic roles are assumed to be cumulative (§2.3): if Rodgers and Hammerstein wrote a musical, and Rodgers and Hart wrote a musical, then Rodgers, Hammerstein, and Hart together wrote *musicals*. (ib) is false because, without supporting context, Champollion's *D* operator only distributes the predicate down to each individual member of the subject, false here because that would require *write a musical* to be true of each artist (whereas with strong contextual support, his *D* operator can distribute to intermediate groupings such as pairs of artists). Champollion interprets (ia)–(ib) as evidence for such a *D* operator. However, I might suggest that (ib) is judged false simply because people prefer non-covarying indefinites over covarying ones (§1.3.4). In that case, the preferred understanding of (ib) is that the three artists co-wrote a single musical — false in the scenario described in (i).

Like *smile*, *see the photo* would also be handled using a meaning postulate guaranteeing that whenever multiple individuals see something, they each do.

(33) Alice and Bob saw the photo.

$see(Alice \oplus Bob, \iota y[photo(y)])$

Meaning postulate: $see(G, y) \rightsquigarrow \forall x[x \in G \rightarrow see(x, y)]$

Meet is handled without a meaning postulate. (Some authors would also apply the group-forming operator \uparrow , but in theory, (2) should already be understood nondistributively without \uparrow — or at least should be compatible with a nondistributive understanding — since there is no meaning postulate making it distributive.)

(34) Alice and Bob met.

$meet(Alice \oplus Bob)$

As for *build a raft*, its distributive understanding is derived using the *D* operator, while its nondistributive understanding is derived when *D* is absent. (As with *meet*, some authors might additionally apply the group-forming \uparrow operator here — more on that shortly.) While meaning postulates cannot be optional, *D* can be present or absent. This ambiguity is used to derive the two distinct ways of understanding (35).

(35) Alice and Bob built a raft.

a. **Distributive:** Alice and Bob *D*(built a raft)

$\forall x[x \in Alice \oplus Bob \rightarrow \exists y[raft(y) \wedge build(x, y)]]$

b. **Nondistributive:** Alice and Bob built a raft

$\exists y[raft(y) \wedge build(Alice \oplus Bob, y)]$

The *D* operator is mainly discussed as a way to handle distributive understandings with covariation, as in (35a), where each person builds a different raft. But *D* must also presumably be used to

derive the distributive understanding of non-covarying predicates such as *open the window* and *lie*. These predicates can be understood nondistributively as well as distributively, so their distributive understanding cannot be captured by a meaning postulate, which cannot be optional. If the only two sources of distributivity are *D* and meaning postulates, then the only alternative is to use *D*. Without *D*, these predicates are presumably understood nondistributively⁶, whereas with *D*, they are distributive.

- (36) Alice and Bob opened the window.
- a. **Distributive:** Alice and Bob *D*(opened the window)
 $\forall x[x \in Alice \oplus Bob \rightarrow open(x, \iota y[window(y)])]$
 - b. **Nondistributive:** Alice and Bob opened the window
 $open(Alice \oplus Bob, \iota y[window(y)])$
- (37) Alice and Bob lied.
- a. **Distributive:** Alice and Bob *D*(lied)
 $\forall x[x \in Alice \oplus Bob \rightarrow lie(x)]$
 - b. **Nondistributive:** Alice and Bob lied
 $lie(Alice \oplus Bob)$

It is useful to consider an intransitive verb such as *lie*, because the *D* operator is generally discussed in the context of multi-word verb phrases. While it is unusual to see *D* applied to a single lexical item, there is no other obvious way to derive the two understandings of *lie* when the only two tools available are *D* and meaning postulates, and a meaning postulate would incorrectly rule

⁶Technically, when *D* is absent, (36b) can actually be understood in both ways (Winter 2000: 5). (36b) simply says that the extension of *open* includes an event with *Alice* \oplus *Bob* as its agent and *the window* as its theme. One way for this to be true is if *Alice* \oplus *Bob* open the window jointly but not individually (nondistributive). But another way for (36b) to be true is if each person opens the window (distributive): on the assumption that verbs and thematic roles are cumulative, if Alice opens the window and Bob opens the window, then *Alice* \oplus *Bob* open *the window* \oplus *the window* — which is just *the window* (on the assumption that something summed with itself is just itself; see Krifka 1992). Thus (36b) is actually compatible with both a distributive understanding and a nondistributive one. The same goes for *lie* in (37b) and any other verb phrase that is closed under sum formation.

out the nondistributive understanding of *lie*.

As for the *lips smiled* example, presumably the meaning postulate requiring *smile* to be distributive no longer applies when its subject is *lips* rather than *humans*.

A note on terminology: Winter 1997 *et seq* and de Vries 2015 use the term ‘P-distributivity’ (short for ‘predicate distributivity’) for the distributive inferences captured by meaning postulates (like the one used for *smile*), on the grounds that these inferences stem purely from world knowledge that the event described by the predicate can only be undertaken individually. They use the term ‘Q-distributivity’ (short for ‘quantificational distributivity’) for the inferences captured by the *D* operator, since *D* quantifies over each member of the subject. Champollion 2010 *et seq* uses the term ‘lexical distributivity’ for distributivity attributed to meaning postulates, and ‘phrasal distributivity’ for distributivity attributed to the *D* operator.⁷

On the two-pronged analysis, the question of which predicates are understood in which ways is split into several parts. Which predicates should be required to be distributive via a meaning postulate? (Presumably *D* is redundant when combined with such predicates.) Which predicates are incompatible with *D*? (Presumably *D* is incompatible with *meet*, at least when its subject is *Alice and Bob*, since individuals cannot *meet* unilaterally.) Which predicates (like *build a raft*) are distributive with *D* and nondistributive otherwise? When does *D* have a hybrid effect — as when it derives the covarying, ‘two-different-photos’ reading of *Alice and Bob saw a photo*, which is still distributive in the absence of *D* given that if two people see something, they each do? Furthermore, if the \uparrow operator is assumed alongside *D*, similar questions arise again: with which predicates is \uparrow redundant, consequential, or incompatible? To explain which predicates go which ways on the two-pronged analysis, these are the questions that must be answered.

⁷In Champollion’s work, the term ‘lexical distributivity’ is also associated with one-word predicates (*smile*), while ‘phrasal distributivity’ is associated with multi-word predicates (*build a raft*). But some one-word predicates (*lie*) can be understood both distributively and nondistributively, presumably handled by *D*; and some multi-word predicates (*see the photo*) are only understood distributively, presumably attributed to a meaning postulate.

3.4.4 Discussion

This section has reviewed three analyses of distributivity from the literature: Landman’s analysis connecting distributivity to plurality; Scha’s meaning postulates; and the widely used two-pronged analysis. While Scha’s meaning postulates cannot capture all the facts, the other two analyses can. I cannot disagree too strongly with them when they capture the same data.

I use the cover analysis (§3.3.2) over these alternatives only because I see it as the most straightforward. While most researchers would agree that a predicate’s potential for distributivity ‘depends on world knowledge’, the cover analysis says that and nothing more. I do not adopt Landman’s ‘distributivity-as-plurality’ analysis because I question its claims about collective / ‘group’ predication (Chapter 2). I do not adopt the widely used two-pronged analysis because I think it complicates the question of which predicates are understood in which ways (involving meaning postulates, D , and perhaps \uparrow , all of which can interact). But these are not knock-down arguments, so readers are welcome to choose an alternative. The claims made in Chapters 4 and 5 are compatible with any analysis that captures the facts.

3.5 Chapter summary

This chapter presents the semantic analysis of distributivity used in this dissertation: a predicate applied to a plural subject is individually true of each cell of a contextually supplied cover of the subject. Inferences about distributivity are framed as inferences about which cover settings to entertain, given what is known about the event described by predicate. Some alternative analyses capture the same data, so the choice between them is not an empirical one.

Like its alternatives, the proposed cover analysis on its own does not make any predictions about which predicates are understood in which ways. Perhaps it is clear why *smile*, *open the window*, *lie*, *build a raft*, and *see the photo* behave the way they do; but there are infinitely more predicates whose distributivity potential remains mysterious. That is why the remainder of this dissertation aims to systematize the behavior of a much wider variety of predicates.

Chapter 4

Verb phrases

This chapter presents the Distributivity Ratings Dataset (Glass & Jiang 2017), in which over 2300 verb phrases are rated for their distributivity potential by online annotators. This dataset allows us to test hypothesized patterns in the behavior of different sorts of verb phrases, so that the underspecified cover semantics from Chapter 3 is complemented by a predictive pragmatic analysis of which cover settings are entertained.

4.1 Introduction

According to the semantics proposed in Chapter 3, *smile*, *meet*, and *open the window* are all represented in the same way: the predicate is true of each cell of a contextually supplied cover of the subject. We draw different inferences from these predicates because we entertain different covers for each one, depending on what we know about these events and how individuals can participate in them. People can only *smile* individually, they can only *meet* multilaterally, they can *open the window* in either way; so *smile* gets a distributive cover, *meet* gets a nondistributive one, *open the window* gets both. The problem is that each predicate is handled on a case-by-case basis, making no clear predictions about the behavior of other predicates. To make the analysis more predictive, the goal is to understand more generally which verb phrases are understood in which ways and why.

4.1.1 Literature motivating the current study

This question is new in that it has not been taken on systematically; but also old, in that it has loomed in the background all along.

It is spotlighted in the work of Dowty 1987. Dowty observes that even nondistributive predicates give rise to inferences that apply distributively to individual members of the subject: (1) is nondistributive, in that only multiple individuals can *gather*; but we also infer that each child was in the hall at the relevant time — a distributive inference, applying to each child.

(1) The children gathered in the hall. adapted Dowty 1987: 99

Dowty calls such inferences ‘distributive sub-entailments’, on the grounds that some sub-component of the predicate’s meaning distributes. (He uses these sub-entailments in the interpretation of the verb phrase modifier *all*.) As for predicates like *smile*, which are understood so that the full predicate distributes to each member of the subject, Dowty views them as a special case of distributive sub-entailments, when the predicate’s sub-entailment is equivalent to the predicate itself.

Dowty then shows that at least some of these sub-entailments are too idiosyncratic to be handled compositionally. For example, in (2), the predicate is understood nondistributively in that it is not individually true of each child; but we also infer that at least 51% of the children *each* voted in favor of the proposal (assuming a majority-based democracy).

(2) The children voted the proposal into effect. adapted Dowty 1987: 99

Similarly, in (3) (which Dowty attributes to personal communication with Rich Thomason), the predicate is understood nondistributively, but we also infer that either exactly one or exactly three of the integers are odd.

(3) These three positive integers sum to thirteen. adapted Dowty 1987: 111

Dowty argues that these inferences do not relate to the logical representations of these sen-

tences, but instead are grounded in extralinguistic facts about democracy and arithmetic, just as the distributive inference associated with *smile* stems from the anatomical fact that people have their own faces. He poses it as a challenge for future work to explain these distributive sub-entailments systematically.

Citing Dowty as inspiration, [Roberts 1987](#) argues that the distributivity of predicates such as *smile* should not be stipulated, because that makes the behavior of *smile* (and similar predicates such as *walk* and *die*) look more arbitrary than it is. Her discussion is worth quoting at length:

‘The fact that a particular lexical item is a group predicate or a distributive predicate doesn’t really need to be specified independently: it follows from the sense of the predicate itself. What does it mean to *gather* or to *disperse*? By virtue of the meaning of such a predicate, its subject must denote a group of individuals (or a mass of some substance), performing in a way peculiar to a group (or mass). [...] What is it to *be a pop star* or to *walk* or to *die*? The actions or states denoted by these verbs can generally only be performed or endured by an individual with a single will and consciousness. It is for this reason that we think of them as distributive. Although it may well be that only atomic individuals are in the extension of such distributive verbs in their strict sense, this follows from our knowledge of what is required for them to be true of an individual.’ ([Roberts 1987](#): 124)

As further evidence against stipulating the distributivity potential of various predicates, [Roberts 1987](#): 124 points out that the behavior of a multi-word predicate depends not only on the main verb, but also on its object and any modifiers (*open the window* can be understood both ways; *open their eyes* only makes sense distributively). So even if every verb were somehow tagged for distributivity, the behavior of full verb phrases would still not follow. But although Roberts is not satisfied with stipulating distributivity, she leaves it for future researchers to offer an alternative.

For Winter and colleagues, inferences about distributivity constitute part of a larger phenomenon: ‘the idea that lexical meanings of predicates may lead to pseudo-quantificational effects’ ([Mador-Haim & Winter 2015](#): 473) — that is, inferences that can be paraphrased using quantificational

language without corresponding to any quantifiers in the logical representation. The inference from *the children slept* to *each child slept*, Winter says, ‘does not need to be regarded as a truth conditional fact about plurals’ (Winter 2001a: 252), but instead arises from our knowledge about sleeping, just as the inference from *the surface is green* to *every part of the surface is green* arises from our knowledge about surfaces and greenness rather than from any covert quantifier. Emphasizing that ‘the link between pseudo-quantification and lexical knowledge is central for semantic theory, an area that is caught between questions about syntactic structure and problems of mental concept modeling’ (Mador-Haim & Winter 2015: 473), he and his coauthors call for more work on the topic:

‘We would like to reiterate the importance that we see for a rigorous theory about the lexicon and the pragmatics of plurals, especially in relation to [...] distributivity [...]. More general and precise theories of these lexical and pragmatic domains will also surely shed more light on the formal semantics of plurality’. (Winter & Scha 2015: 35)

As long as researchers have known about distributivity, they have known that it is fundamentally shaped by world knowledge, and have periodically challenged future researchers to make this idea predictive.¹ But the challenge is still outstanding.

4.1.2 Where the current work fits in

To make progress, this chapter presents the first large-scale study of the distributivity potential of verb phrases (§4.2; Glass & Jiang 2017), using data which I collected along with Nanjiang Jiang, a summer intern at Stanford’s Center for the Study of Language and Information. The dataset provides quantitative ratings from online participants for questions of the form (4a) and (4b) for

¹Similar issues also surface among reciprocal predicates, as in *the children know each other* (where each child knows every other child) or *the plates are stacked on top of each other* (where each plate except the bottom one is stacked directly on top of one other plate); Dalrymple et al. 1994; Winter 1996; Dalrymple et al. 1998; Winter 2001b, Poortman et al. 2018, Winter 2018. These authors propose that a reciprocal sentence expresses the logically ‘strongest’ (Dalrymple et al) or most ‘typical’ (Poortman et al. 2018, Winter 2018) truth conditions compatible with what is known about the predicate, calling on world knowledge and lexical semantics to constrain the compositional meaning of such sentences and raising questions about how exactly this ‘strongest’ or most ‘typical’ meaning is calculated.

over 2300 verbs, categorized by meaning using the system of [Levin 1993](#). (Transitive verbs were given singular, indefinite objects following a process described in §4.2.1.)

- (4) Naomi and Jeff {smiled, opened a window, ...}.
- a. Does it follow that Naomi and Jeff *each* {smiled, opened a window, ...}?
- definitely no maybe no not sure maybe yes definitely yes
- b. Could it be that Naomi and Jeff didn't technically *each* {smile, open a window, ...}, because they did so *together*?
- definitely no maybe no not sure maybe yes definitely yes

This dataset makes it possible to test hypotheses about the behavior of various types of predicates. In particular (§4.3), we might expect that the distributivity potential of a given verb phrase should align with certain lexical semantic properties — whether it involves a transitive verb or an intransitive one; whether it describes an event carried out by an individual body or mind; whether it describes an inherently multilateral event; whether it is causative; whether it has an incremental object in the sense of [Tenny 1987](#), [Krifka 1989](#), and [Dowty 1991](#). If we assume that these properties of verb phrases map onto conceptually and inferentially significant aspects of the events they describe, then we predict that these properties should help to determine their distributivity potential.

§4.2 presents the dataset. §4.3 motivates a series of hypotheses about the distributivity potential of various types of predicates (repeated from §1.4), and tests these hypotheses empirically:

- **TRANSITIVE / INTRANSITIVE HYPOTHESIS:** Predicates built from many intransitive verbs (*smile*) can only be distributive, while those built from many transitive verbs (*open the window*) can be understood nondistributively as well as distributively ([Link 1983](#), [Glass 2017](#)).
- **BODY / MIND HYPOTHESIS:** Predicates describing bodily or mental actions (*smile, jump, meditate, swallow a pill, see a photo, like a book*) are understood distributively, given that individuals have their own bodies and minds and so can only carry out these events individu-

ally.

- **MULTILATERAL HYPOTHESIS:** Predicates describing inherently multilateral actions (*meet*) are understood nondistributively, given that individuals cannot carry out such actions alone.
- **CAUSATIVE HYPOTHESIS:** Predicates built from causative verbs (describing an event where the subject causes the object to change, such as *open the window*) can be understood nondistributively (as well as, perhaps distributively, depending on definiteness and repeatability), given that the nature of causation allows that multiple individuals' contributions may be jointly sufficient but individually insufficient to cause a result.
- **INCREMENTAL HYPOTHESIS:** Predicates with incremental objects (those whose parts correspond to parts of the event described by the predicate, such as *eat the pizza*) can be understood nondistributively (as well as, perhaps, distributively), given that multiple individuals might each carry out the event described by the verb on a different portion of an incremental object (eat a different part of a pizza), only jointly adding up to (eating) the whole thing.

4.2 Distributivity Ratings Dataset

The main empirical contribution of this work is the Distributivity Ratings Dataset (Glass & Jiang 2017), which provides quantitative ratings for the distributivity potential of over 2300 predicates constructed from the verbs of Levin 1993.

Levin 1993 organizes over three thousand English verbs into nearly two hundred classes of verbs with similar meanings (that is, describing similar types of events).² For example:

- (5) **'Put' verbs** arrange, immerse, install, lodge, mount, place, position, put, set, situate, sling, stash, stow Levin 1993: Chapter 9

²Levin also organizes the verbs into 'alternation classes' based on their argument structure and syntactic behavior, arguing that verbs with similar meanings pattern together syntactically; but I only use the meaning-based 'verb classes' of her Chapters 9 to 57, not the syntactic 'alternation classes' of Chapters 1 through 8.

- (6) **‘Amuse’ verbs** abash, affect, afflict, affront, aggravate, agitate, agonize, alarm, alienate, amaze, amuse, anger, annoy, antagonize, appall, appease, arouse, assuage, astonish, astound, awe, baffle, beguile, bewilder, bewitch, boggle, bore, bother, bug, calm, captivate, chagrin, charm, cheer, chill, comfort, concern, confound, confuse, console, content [...and many more] Levin 1993: Chapter 31

The Levin classification serves as the starting point for this study firstly because, by listing many of the verbs of English, it provides the material to study verb phrases at a large scale. Moreover, by grouping verbs into classes based on the sorts of events they describe, it offers a way to test the idea that a predicate’s distributivity potential is shaped by that event. One would expect that verbs describing similar sorts of events (those within a Levin class, or those within related Levin classes) should pattern together with respect to distributivity.

The materials for the online study were built using the Levin verbs. These verbs had to be placed into sentences, which were generated automatically. Each sentence was given as its subject a conjunction of two names, chosen randomly from a list (*Veronika, Ian, Luke, Olivia . . .*). Because the stimulus sentences had names as the subject, all verbs that do not make sense applied to humans were excluded — for example, weather verbs (*rain, drizzle*), verbs describing animal reproduction (*calve*), non-human spatial verbs (*border*), and so on. Because the stimulus sentences strictly follow a ‘subject-verb-object’ format, I also excluded verbs requiring prepositional phrases (*put a book on the table*), or elements other than noun phrases as complements (*decree that smoking is illegal, masquerade as an official, keep swimming*).

For intransitive verbs, sentences were generated following the form of (7).

- (7) Name1 and Name2 verbed.
Example: Veronika and Ian giggled.

For transitive verbs, sentences followed the form of (8).

- (8) Name1 and Name2 verbed an object.

Example: Luke and Olivia wrote a book.

To generate these sentences, it was necessary to find an appropriate object for every transitive verb.

4.2.1 Choosing objects for transitives

As mentioned above (§1.3.3), the object of a transitive verb plays an important role in shaping the distributivity potential of the full verb phrase. Indefinite objects can ‘covary’, while definite objects cannot (which interacts with the issue of whether the action described by the verb can be repeated on the same object; §1.3.3). Plural objects systematically create the potential for a nondistributive ‘cumulative’ understanding (Chapter 2) — if Alice and Bob saw *two photos*, perhaps they each saw one, adding up to two photos between them (nondistributive, because *see two photos* is not individually true of Alice or of Bob).

For the Distributivity Ratings Dataset, objects had to satisfy two criteria: they had to be indefinite, in order to abstract away from the issue of whether the action described by the verb can be repeated on the same object; and they had to be singular, to avoid the potential for cumulativeness discussed in Chapter 2. That way, sentences built from transitive verbs are all modeled on the frame in (8).

Beyond the grammatical features of being definite / indefinite or singular / plural, the distributivity potential of a verb phrase is also influenced by the referent of its object (§1.3.3). *Open their eyes* (or, using a singular indefinite object, *open an eye*) is understood distributively given that people have their own eyes. *Open a vault* is likely to differ from *open a soda* given the sizes of these objects and the difficulty of opening each one.

It therefore seems important to choose objects for verbs in the Distributivity Ratings Dataset using a method that systematically controls for these issues. Particularly if the focus of the study is *verbs*, we do not want the choice of object to confound the data. But it is not obvious what method would control for such confounds. We certainly cannot give every verb the same object (*open a*

window vs. #*eat a window*); and a generic object such as *thing* would be unnatural.

In the era of ‘big data’, it may seem like the answer is to simply choose the most frequent object for each verb from corpus data. But such an off-the-shelf method becomes messy. Some verbs would be given body-part objects (which are often strange as singular indefinites; *shake a head*, *wrinkle a nose*); container or unit nouns (*cook a minute*, *mince a tablespoon*); objects that are part of frozen or metaphorical expressions (*keep an eye*, *abhor a vacuum*); relational nouns that sound strange out of context (*find a way*); or objects that do not make sense in the context of the Levin class within which the verb is classified (*snap a photo* when *snap* is categorized as a change-of-state verb). Corpus data is indispensable for finding naturalistically motivated objects; but it cannot be used indiscriminately.

As a compromise, my strategy was to generate for each verb a set of candidate objects from corpus data (specifically, the 30 most frequent nouns to occur within 5 words following that verb in the part-of-speech-tagged Spoken section of the Corpus of Contemporary American English; Davies 2008),³ and then to hand-select the ‘best’ object from among these candidates, based on a list of criteria that I developed:

1. The object has to make sense as a singular indefinite in a sentence of the form in (8) (Name1 and Name2 verbed an object). Therefore, less-relational nouns are preferred over more-relational nouns (*find a solution* over *find a way*). Similarly, nouns that are more natural as indefinites are preferred (*view a videotape* over *view a world*).
2. The object has to be construable as a count noun (*melt a chocolate* over *melt an ice*).
3. The object has to make sense within the Levin class in which the verb is classified. When *snap* is classified as a change-of-state verb, *a twig* is preferred over *a picture*. When *hang* is classified as a ‘put’ verb, *a picture* is preferred over *a prisoner*.

³We searched for all nouns that occurred within five words following the verb — not just the singular indefinite ones. That way, the list of candidate objects for each verb included ones that were used as definites, plurals, possessive DPs, things modified by adjectives, and so on: if the string *entertained their young children* appeared in the Spoken CoCA data, then *child* could appear among the candidate objects for *entertain*. This methodology offered more data than if we had only considered unmodified singular indefinite objects.

4. When possible, the object should be concrete rather than abstract: *squash a bug* is preferred over *squash a hope*.
5. The object should not be part of an idiomatic or metaphorical expression: *dodge a question* is preferred over *dodge a bullet* (used non-literally to mean ‘escaping a bad situation’).
6. The object should not be a negative polarity item (*lift a finger*).
7. The object should not be a body part (*slit a skirt* over *slit a throat*), because the same predicate might only be understood distributively with a body-part object when it could be understood nondistributively with a different object (*open an eye* versus *open a window*). The only exception is Levin’s class of ‘Verbs Involving the Body’ (*skin a knee*, *twist an ankle*, and so on), which were given body-part objects.
8. When possible, the object should not profile specific demographic groups (*persecute a minority* is preferred over *persecute a Christian / Jew*), nor should it create an excessively violent sentence. Among the verbs describing violent actions, there were still some upsetting sentences (*suffocate an infant*, *drown a child*, and so on); I added a note in the introduction so participants would not be unpleasantly surprised.
9. While optimizing all these constraints, more-frequent objects are favored over less-frequent ones.
10. If none of the 30 candidate objects make sense (or if fewer than 30 were generated because the verb is infrequent), the example sentences given in the *Oxford Advanced Learners’ Dictionary* (Hornby et al. 1995) are consulted; if no suitable objects are found there either, then the verb is excluded.

As an example, for *lift*, the object *a boat* was chosen among the candidates in (9).

- (9) weight, arm, head, ban, leg, sanction, spirit, foot, hand, right, hip, embargo, finger, body, time, ticket, eye, lid, chest, bar, people, pound, **boat**, heel, torso, restriction, object, chin,

knee, export

For *stow*, *a canoe* was chosen among the candidates in (10):

(10) *stow*

gear, item, antique, rod, dinghy, ugo, space, **canoe**, resident, nut, goggle, terrain, mission, chalk, attitude, supplement, deadline, fishing, mirror, saving, array, wakeboards, hook, bit, half-liter, bag, recruiter, point, paddle, rubber

For *tile*, *a bathroom* was chosen among the candidates in (11):

(11) *tile*

duty, restriction, tournament, dotted-dashed, overlap, human, flavor, trust, **bathroom**, area, sky, brunt, print, proposal, variability, clean, alhambra, villager, park, sophism, floor, line, present, man, keyword, universe, ncaa, backsplash, roman, floors

This blend of bottom-up and top-down methods yields objects that are both naturalistically motivated and controlled for various confounds.⁴

4.2.2 Study design

Verb phrases that were tested In Levin’s data, the same verb often appears in multiple classes. *Cackle* is both an ‘animal sound’ verb and a ‘nonverbal expression’ verb. *Arrange* is a ‘put’ verb and a ‘build’ verb. *Beat* is a ‘sound’ verb, a ‘hit’ verb, and a ‘knead’ verb. If the verb is intransitive, then it is tested in the same way regardless of its Levin class: it is placed into a sentence of the form ‘Name1 and Name2 verbed’. But if the verb is transitive, then it may have a different object in different Levin classes. When *beat* is a ‘sound’ verb and a ‘hit’ verb, its object is *a drum*; but when it is a ‘knead’ verb, its object is *an egg*. Therefore, *beat* is tested in two different formats (‘Name1 and Name2 beat a drum’ and ‘Name1 and Name2 beat an egg’) across three different Levin classes.

⁴Thanks to Chris Potts for discussion.

The data were compressed into the set of unique predicates. Any intransitive verb (such as *cackle*) appears in these de-duplicated data only once, with a list of its Levin classes. Any transitive verb appears in the data once for each distinct object with which it was tested: *beat* is listed once with the object *a drum* (spanning both the ‘sound’ and ‘hit’ classes), and once with the object *an egg* (the ‘knead’ class). De-duplicated in this way, we are left with 2338 unique verb phrases (1667 transitive, 671 intransitive).

Stimuli Next, judgments were elicited from online participants about the distributivity potential of these predicates. As previewed above, participants answered questions of the form in (12) for each one. The five answer choices (*definitely no*, *maybe no*, *not sure*, *maybe yes*, *definitely yes*) were recorded on a 1–5 Likert scale.

(12) Naomi and Jeff {smiled, opened a window, ...}.

a. Does it follow that Naomi and Jeff *each* {smiled, opened a window, ...}?

definitely no maybe no not sure maybe yes definitely yes

b. Could it be that Naomi and Jeff didn’t technically *each* {smile, open a window, ...}, because they did so *together*?

definitely no maybe no not sure maybe yes definitely yes

It may seem strange that the two questions (12a) ‘each’ and (12b) ‘together’ are not symmetrical; (12a) ‘each’ asks about what ‘follows’ while (12b) ‘together’ asks about what ‘could be’. I chose this wording because I wanted (12a) ‘each’ and (12b) ‘together’ to probe from two different angles at the same issue: whether the predicate is *only* understood distributively, or whether it can *also* be understood nondistributively. (As expected, responses to the two questions are highly negatively correlated — see below.)

In general, most predicates describe events that an individual could plausibly undertake individually (*smile*, *open the / a window*). A finite number of predicates (*meet*, *gather*) describe inherently

multilateral actions; but they are the exception rather than the rule. Since most predicates describe events that can be carried out by individuals, most predicates can be understood distributively when applied to a plural. The only further exception here is predicates with definite objects describing events that cannot be repeated on the same object (*break the vase* cannot be understood distributively, given that the same vase cannot be broken more than once; §1.3.3). But when we restrict our attention to predicates with indefinite objects, then apart from the *meet*-type predicates, predicates in general can be understood distributively when applied to a plural.

As a result, it is most informative to investigate which predicates have an available nondistributive understanding in addition to a distributive one. Some predicates behave like *smile* in only making sense distributively; others behave like *open the / a window* in also allowing a nondistributive understanding. The questions in (12a)–(12b) are designed to distinguish the *smile* type from the *open the / a window* type. For *smile*-type predicates, the response for (12a) ‘each’ should be high while the response for (12b) ‘together’ should be low. For *open the / a window*-type predicates, the response for (12a) ‘each’ should be low while the response for (12b) ‘together’ should be high. (For a *meet*-type predicate, the response for (12a) ‘each’ should be even lower.) These questions therefore divide predicates in an informative manner.

Counterfactually, if (12a) ‘each’ asked about what ‘could be’ as opposed to what ‘follows’, then nearly every predicate other than the *meet*-type ones would be predicted to have a high rating, making that question uninformative. Similarly, if (12b) ‘together’ asked about what ‘follows’ rather than what ‘could be’, then nearly every predicate other than the *meet*-type ones would be predicted to have a low rating, making that question uninformative too. In contrast, the actual questions (12a) ‘each’ and (12b) ‘together’ are designed to divide predicates in the most informative way — into those that can only be understood distributively, like *smile*, and those that can also be understood nondistributively, like *open the / a window*.

It is also important that the two opposing perspectives in (12a) ‘each’ and (12b) ‘together’ force participants to consider both distributive and nondistributive understandings — which is why two questions were used as opposed to just one. Otherwise, without having to imagine both possibilities,

I feared that participants would be too generous in allowing that two people who participate in a particular event ‘each’ carry out that event.

One might also worry about the word *together* in the ‘together’ question (12b). *Together* is notoriously polysemous (Lasersohn 1988 / Lasersohn 1990b, Lasersohn 1990a, Schwarzschild 1992, Lasersohn 1998a, Moltmann 2004, Syrett & Musolino 2013, Syrett & Musolino 2016); if two individuals carry out a predicate *together*, perhaps they were in the same place at the same time, and / or they coordinated their actions. In that case, the predicate might still be individually true of each of them (distributive; *Alice and Bob smiled together*). Or perhaps *together* indicates that the predicate is not true of each individual, but only true of the two of them *together* (nondistributive; *Alice and Bob opened the window together*). If *together* is understood in its ‘proximity’ or ‘coordinated action’ senses rather than its sense as a ‘nondistributivity marker’ (Schwarzschild 1992), then the question (12b) would be confounded. But I believe that the surrounding context — ‘didn’t technically *each* VP, because they did so *together*’ — helps to disambiguate. And indeed, as illustrated below, the participants’ responses largely indicate that they understood it as a nondistributivity marker, as intended.

Finally, the participants did not receive any training on how to interpret the questions. Their uncertainty may have led them to choose the less-committal answer choices ‘2=maybe no’ and ‘4=maybe yes’ over the poles (‘1=definitely no’ and ‘5=definitely yes’). (See the Appendix 6.3 for a followup experiment where participants were trained on how to interpret the questions, and were much more willing to choose the ends of the scale.) Although the effect sizes might have been larger if participants had felt more certain in their answers, their responses convey that they interpreted the questions as intended.

Data collection We used web developer tools (JavaScript, jquery, HTML / CSS) to create an online survey in which participants encountered 40 questions of the form in (12) (with two subquestions per question, resulting in 80 datapoints per participant). Each participant therefore only saw a randomized subset of 40 predicates from the 2338 unique predicates that we tested. Questions appeared in a random order. There were no fillers because there was no controlled manipulation

that fillers would serve to disguise. An optional checkbox was added to each question allowing the participant to indicate that they were unfamiliar with the given verb (because some of the verbs, such as *pip* and *carom*, were quite rare); if a participant checked that box, their responses for that predicate were ignored.

Instructions

In this experiment, you will encounter sentences of English and answer questions about them. The sentences will describe an action done by multiple people, and you'll answer questions about what you think each individual person did as part of that action.

These sentences were generated from naturally occurring data, so sometimes they might not totally make sense out of context, but you should do your best to understand them. Don't spend too long thinking and re-thinking each question; just go with what feels right to you. We regret that these sentences may occasionally describe events that are somewhat violent, but our study requires us to sample a variety of event types.

Thanks again for helping us with our research.

Start experiment

Figure 4.1: Screenshot of the instructions page.

Participants were recruited using Amazon's Mechanical Turk service. They all used United States I.P. addresses. They were paid \$2.00 for what amounted to approximately 9 to 15 minutes of work (excluding severe outliers), so that the slower participants earned \$8 / hour and the faster participants earned more. Participants were asked to report their native language at the end of the experiment after being advised that they would be paid regardless of their answer; data were only analyzed from those who reported that their native language was English.

The goal was to collect three observations for each of the 2338 unique predicates. We did not have a way to keep track of how many observations had been recorded for each predicate; instead,

1 / 40

Adam and Lee abandoned a ship.

Does it follow that Lee and Adam *each* abandoned a ship?

definitely no
 maybe no
 not sure
 maybe yes
 definitely yes

Could it be that Lee and Adam didn't technically *each* abandon a ship, because they did so *together*?

definitely no
 maybe no
 not sure
 maybe yes
 definitely yes

I don't know what "abandon" means.

Continue

Figure 4.2: Screenshot of an item from the experiment.

we hoped that with enough participants, we would eventually get three observations per predicate. This methodology was perhaps not the most efficient, because some predicates were ultimately seen over ten times. We initially ran 270 participants (more than enough for each predicate to be seen 3 times). But given our setup, some predicates were seen more times than needed, while others were seen fewer than 3 times. To get at least three observations per predicate, we ran 58 additional participants, using only the predicates that had not received three ratings initially. After excluding three non-native English speakers, and removing the 484 observations for which the participant said they did not know the verb (for verbs such as *confabulate*, *jeep*, *scrawk*, *ululate*, and *scud*), we ended up with 325 participants and 12,515 responses for the questions represented in (12); with two subquestions per question ((12a) ‘each’ and (12b) ‘together’), there are 25,030 datapoints in all.

4.2.3 Results

The results can be formatted in two different ways. In one format (Table 4.1), each row gives a predicate along with its average ratings for both the (a) ‘each’ question and the (b) ‘together’ question. Table 4.1 shows the average ratings for the predicates discussed in Chapter 3 (except, unfortunately, for *lie*, because the ‘mislead’ meaning of *lie* is not classified by Levin; note also that the object of *build* is *a house*, not *a raft*).

verb	object	‘each’ avg.	‘together’ avg.	Levin class
smile	n/a	4.57	1.57	‘nonverbal expression’ verbs
meet	n/a	2.67	5.00	‘meet’ verbs
open	a window	3.60	3.80	‘other change-of-state’ verbs
build	a house	2.50	4.50	‘build’ verbs
see	a photo	4.50	2.00	‘see’ verbs
(...)	(...)	(...)	(...)	(...)

Table 4.1: Average ratings for both ‘each’ and ‘together’ questions for each predicate.

The full csv file also includes a column listing the scores given by each participant who rated that predicate (elided here for ease of reading). This format obscures the contributions of each participant to highlight the aggregate behavior of each predicate.

In the other format (Table 4.2), each row gives an individual participant’s ratings for a given predicate. (The same predicate is thus listed in multiple rows, once for each participant who encountered it.) Table 4.2 illustrates the first three participants in our data, along with some of the predicates that each of them saw.

This format obscures the aggregate ratings for each predicate to expose the contribution of each participant. It facilitates statistical tests that include ‘participant’ as a random effect, allowing us to factor out unexplained differences between individual participants (discussed further below).

Table 4.3 shows the number of responses in each of the five response categories for both the ‘each’ question and the ‘together’ question. For both questions, ‘maybe yes’ is the most common answer, and ‘not sure’ is the least common.

All of these data, in both formats (and the R code used for the analysis), are publicly available

ParticID	verb	object	'each'	'together'	Levin class
Subj1	crack	an egg	2	4	'break' verbs
Subj1	cackle	n/a	4	2	'manner of speaking/animal sound/nonverbal expression' verbs
(...)	(...)	(...)	(...)	(...)	(...)
Subj2	steady	a canoe	4	5	'other change-of-state' verbs
Subj2	resent	an intrusion	4	4	'admire' verbs
(...)	(...)	(...)	(...)	(...)	(...)
Subj3	wheeze	n/a	4	2	'manner of speaking/hiccup/sound emission' verbs
Subj3	bend	a wire	2	4	'knead/bend' verbs

Table 4.2: Each participant's ratings for both the 'each' and 'together' questions for each predicate they encountered.

Response	Total for (a) 'each'	Total for (b) 'together'
1 (definitely no)	1055 (8.4%)	1750 (14.0%)
2 (maybe no)	2167 (17.3%)	1708 (13.6%)
3 (not sure)	788 (6.3%)	874 (7.0%)
4 (maybe yes)	4683 (37.4%)	5189 (41.5%)
5 (definitely yes)	3820 (30.5%)	2992 (23.9%)

Table 4.3: Number and percentage of responses in each of the five response categories for both the 'each' and 'together' questions.

through the Open Science Framework at <https://osf.io/8953e/>.

Mixed-effects linear regression and a sanity check The format of Table 4.2, where each line represents one participant's encounter with one verb phrase, is suitable for the statistical test I use throughout this chapter: mixed-effects linear regression, conducted using the `lme4` package of R (Bates et al. 2015b; see B. Winter 2013 for a tutorial on such models as they are used in linguistics, and Gelman & Hill 2007 for helpful general discussion.)

To demonstrate mixed-effects linear regression, I first perform a 'sanity check', testing whether a predicate's rating for the 'each' question is related to its rating for the 'together' question. A high rating for 'each' indicates that the predicate is understood distributively, while a high rating for 'together' shows that it can be understood nondistributively. So we expect that the higher a predicate's rating for the 'each' question (the more it wants to be understood distributively), the

lower its rating should be for the ‘together’ question (the less it makes sense nondistributively), and vice versa.

A mixed-effects linear regression predicts a continuous dependent variable on the basis of one or more (continuous or categorical) independent variables in a way that *factors out* other random contributions to this dependent variable that are unrelated to the hypothesis being tested. Here, we are predicting a participant’s response to the ‘each’ question (the dependent variable) on the basis of their response to the ‘together’ question (the independent variable, also known as the ‘fixed effect’).

A participant’s ‘each’ rating for a given predicate does not just depend on their ‘together’ rating for that predicate (the ‘fixed effect’), but also on how the specific participant tends to use the ratings scale (a ‘random effect’), and also on the specific verb phrase (another ‘random effect’). To test the prediction that a predicate’s ‘each’ rating is related to its corresponding ‘together’ rating, it is important to factor out the ‘random effects’ of differences between individual participants or verb phrases (mathematically, the model allows the intercept in the linear regression to vary with each participant and each verb phrase). Such a mixed-effects structure makes use of all the available information — that the same participant rated multiple different predicates, and that the same predicate was rated by multiple different participants — and uses this information to help explain the variance in distributivity ratings. In this way, it is a ‘conservative’ model, unlikely to find a spurious effect.

It is important to remember that experimental participants chose among five responses — ‘definitely no’, ‘maybe no’, ‘not sure’, ‘maybe yes’, and ‘definitely yes’ — which are mapped to a one-to-five Likert scale (§4.2.2) for the statistical analysis. In other words, I am treating what is technically an ordered categorical variable as a linear, continuous one: assuming that the difference between ‘definitely no’ and ‘maybe no’ is equal to the difference between ‘maybe no’ and ‘not sure’, just as the difference between 1 and 2 is equal to that between 2 and 3. This way of handling Likert data is extremely common and arguably justified in work on psychology and linguistics (see e.g. [Carifio & Perla 2007](#), [Brown 2011](#)).

I used the `lme4` package of R ([Bates et al. 2015b](#)) to run a mixed-effects linear regression using a

predicate’s rating for ‘together’ to predict its rating for ‘each’ (giving each individual participant and predicate a random intercept, meaning that the model attributes some of the variance to unexplained differences between participants and predicates).

R command for the model testing the relationship between ‘each’ and ‘together’ questions

```
lmer(each_rating ~ together_rating
      + (1 | SubjId)
      + (1 | full_pred),
      data = d)
```

Indeed, a predicate’s ‘together’ rating is highly predictive of its ‘each’ rating; for every 1-point increase in its average for ‘together’, its rating for ‘each’ is predicted to decrease by 0.54 points (a highly significant effect at $p < 0.0001$), so that a predicate with a rating of 2 for the (b) ‘together’ question is predicted to have an ‘each’ rating of 4.45, and a predicate with a rating of 5 for the ‘together’ question is predicted to have an ‘each’ rating of 2.83. (Even though participants were restricted to choosing among five options mapped to integers, the statistical model predicts decimals because it treats the scale as linear.) In sum, the two questions are strongly negatively correlated, even if not perfectly so (even if a one-point increase in the ‘together’ rating does not entail a matching one-point increase in its ‘each’ rating). I conclude that the ‘each’ and ‘together’ questions explore the same issue from different angles, as intended.

4.3 Motivating and testing hypotheses

Having introduced the dataset and the statistics used to analyze it, the next step is to test hypotheses about the way different types of predicates should behave.

4.3.1 Full models including all predictors

All of the results reported below are drawn from two full models — one for the ‘each’ question, one for the ‘together’ question — including all of the independent variables hypothesized to predict a predicate’s distributivity potential. As elaborated below, these independent variables include:

1. whether the verb is transitive or intransitive (§4.3.2)
2. whether or not the verb describes an action carried out by an individual body or mind (§4.3.3)
3. whether or not the verb describes an inherently multilateral action (§4.3.4)
4. whether or not the verb is causative (§4.3.5)
5. whether or not the object can be construed as incremental (§4.3.6)
6. ... and (in some models but not others) some interactions:
 - (a) interaction between ‘transitive / intransitive’ and ‘body / mind’
 - (b) interaction between ‘body / mind’ and ‘incremental’
 - (c) interaction between ‘causative’ and ‘incremental’

One model predicts a predicate’s ‘each’ rating as a function of all these independent variables, allowing intercepts to vary for both participants and predicates.⁵ Another model predicts a predicate’s ‘together’ rating as a function of the same independent variables, again allowing intercepts to vary for participants and predicates.⁶ Whereas the model used above to illustrate mixed-effects linear regression used a single, ‘continuous’ predictor (the predicate’s ‘together’ rating, treated as

⁵There is a debate in the literature about how to use random effects: whether the model should always use the maximal number of parameters justified by the study design (Barr et al. 2013), or whether one should decide on a case-by-case basis which random effects actually contribute to the model (Bates et al. 2015a). In the spirit of Barr et al. 2013, I tried to run models for both the ‘each’ question and ‘together’ question using all the fixed effects in 1–5 alongside the ‘maximal’ random effects structure (allowing random slopes for each participant depending on each fixed effect, meaning that the model would allow each participant to not just use the ratings scale differently, but also to respond differently to each fixed effect). But these models fail to converge, meaning that there is not enough data to estimate all of these different parameters. Some models converge when subsets of the maximal possible random effects are used: for example, when each participant’s slope is allowed to vary depending on whether the verb is transitive or intransitive (but not depending on whether it is a body / mind verb, multilateral, causative, or incremental); in those cases, all the results reported below remain significant. Because models with the full random effects structure do not converge, I let only the intercept, not the slopes, vary for each ‘participant’ and ‘predicate’, using more parsimonious models in the spirit of Bates et al. 2015a.

⁶Note that variance attributed to these random effects (0.39 for each participant, 0.12 for each predicate in the ‘each’ model; 0.32 and 0.17 for the ‘together’ model) is small in comparison to the residual variance (0.99 for the ‘each’ model, 1.15 for the ‘together’ model), meaning that the unexplained differences between individual participants and predicates have a relatively small effect on distributivity ratings.

continuous), these models use multiple, binary categorical predictors: whether the verb is transitive or intransitive; whether the predicate is tagged as a body / mind predicate or not; and so on (1–5).

By including all of these fixed effects (1–5) at once, these combined models allow us to isolate the effect of each independent variable, which is important because they overlap (Table 4.5): for example, 112 of the 1667 transitive verbs are body / mind verbs (7%); and 945 of the 1667 transitive verbs are causative (57%).

	trans	intrans	body/mind	multilateral	causative	incr	total
trans	1667 (100%)	0	112 (7%)	0	945 (57%)	201 (12%)	1667
intrans	0	671 (100%)	364 (54%)	91 (14%)	0	0	671
body/mind	112 (24%)	364 (76%)	476 (100%)	0	0	21 (4%)	476
multilateral	0	91 (100%)	0	91 (100%)	0	0	91
causative	945 (100%)	0	0	0	945 (100%)	55 (6%)	945
incr	201 (100%)	0	21 (10%)	0	55 (27%)	201 (100%)	201

Table 4.4: Number of predicates in each category, and overlap between the categories.

For example (discussed further below), most body / mind verbs are intransitive (in fact, 76% of them are). A model which just used one of these independent variables or the other would conflate the effects of each one: if intransitives are found to differ from transitives, for example, we wouldn't know if this effect is driven only by the body / mind intransitives. In contrast, a model which includes both independent variables isolates the effect of each; if each one is significant, it is predictive independent of the other. Similarly, all causatives as defined here are transitive. A model which just used one independent variable or the other (transitive or causative) would blend these effects together: if transitives differ from intransitives, we wouldn't know if this effect is driven only by causative transitives (which in fact are 57% of all transitives; see Table 4.5); if causatives differ from non-causatives, we wouldn't know if this effect is driven only by the fact that all causatives are transitive. But a model including both independent variables reveals the effect of being causative above and beyond being transitive and vice versa. Furthermore, all causatives are transitive and most (76%) of the body / mind verbs are intransitive (Table 4.5), so only a model using all three of these independent variables (transitive / intransitive, causative / non-causative, and body-mind /

non-body / mind) can disentangle these effects.⁷

In what follows, I show that each of the hypothesized independent variables in (1)–(5) significantly predicts the distributivity potential of a verb phrase — both its ‘each’ rating and ‘together’ rating. Since these findings are drawn from a combined model, we can be sure that each effect persists independently of the others.

Finally, I ran the combined models both with and without some interaction terms (6a)–(6c). I did not have a hypothesis about these interactions, but I tested them for completeness. As discussed below (see Table 4.5), most body / mind verbs are intransitive, but some are transitive; so I allowed the model to make different predictions for verbs that were *both* transitive *and* body / mind verbs (*swallow a pill*; *see a photo*). Similarly, some causative predicates can have incremental objects (*cube a zucchini*: the zucchini is causally affected, in that it is cut into cubes, but also potentially incrementally affected, in that each part of it may be cubed in sequence), so I allowed the model to make different predictions for predicates falling into both of these categories. And some incremental-object predicates involve body / mind verbs (*eat a pizza*: the pizza is incrementally affected, and *eating* requires an individual body and digestive system), so I allowed the model to make different predictions for these too. (No other interactions were justified because no other categories

⁷Since the independent variables overlap to some degree, one might be concerned about multicollinearity, meaning that the independent variables are too tightly correlated, which can cause the model to inaccurately estimate the effect of these independent variables. To quantify the collinearity of these independent variables, I used code written by Florian Jaeger (<https://hlplab.wordpress.com/2011/02/24/diagnosing-collinearity-in-lme4/>) to conduct VIF (Variance Inflation Factor) tests on both the ‘each’ model and ‘together’ model (using the fixed effects in 1–5 but no interactions, and allowing intercepts to vary for participants and predicates). The results:

- (i) VIF score for (a) ‘each’ question; (b) ‘together’ question:
 - a. Transitive/intransitive: 2.09; 2.08
 - b. Body/mind: 1.69; 1.69
 - c. Multilateral: 1.25; 1.25
 - d. Causative: 1.50; 1.50
 - e. Incremental theme: 1.09; 1.09

In general, a VIF score below 5 (and certainly below 2.5) indicates no cause for concern. So I conclude that even though the independent variables overlap to some degree, their multicollinearity is not a problem for the statistical analysis. In any case, one of the biggest problems with multicollinearity is that it obscures the significance of correlated independent variables; but since all of these independent variables are statistically significant here, that problem is moot.

cross-cut each other the way these do.⁸) An Akaike Information Criterion (AIC; Akaike 1974) comparison shows that the ‘best’ (lowest-AIC; most predictive and parsimonious) model for the ‘each’ question includes 1–5 and 6a, but not 6b–6c; while the ‘best’ (lowest-AIC) model for the ‘together’ question includes only 1–5 and no interactions. The statistics reported below are taken from these ‘best’ models according to the AIC; but in any case, all of the main effects reported below remain significant whether these interactions are included or not. In general, the results reported below persist regardless of how one chooses to run the statistics, indicating that they are quite robust.

In the tables given below, I report the predicted ‘each’ or ‘together’ ratings, β coefficients, standard errors, degrees of freedom, t statistics, and significance levels (p) for various models. To review these terms:

- **Predicted ‘each’ (or ‘together’) rating:** Predicted rating (along a 1-5 Likert scale treated as continuous) for the ‘each’ (or ‘together’) question for a predicate of the relevant type. Calculated by adding or subtracting the relevant β coefficients from the intercept; for example (Table 4.5), a regular intransitive is predicted at 4.09; a transitive is predicted at -0.58 points less than that (=3.51).
- **β coefficient:** The number added or subtracted from the intercept to predict the ‘each’ or ‘together’ rating for the relevant type of predicate. For example (Table 4.5), the β score for a transitive verb is -0.58, which means we subtract that from the intercept (4.09) to get the model’s predicted ‘each’ rating for a transitive verb.
- **Standard Error (SE):** A measurement of the accuracy of the model’s predictions, defined as $\sqrt{\sum(Y - Y')^2/N}$, where Y is an actual predicate’s ‘each’ (or ‘together’) rating, Y' is its predicted rating according to the model, and N is the number of pairs of Y, Y' . The closer the predicted values Y' are to the actual values Y , the lower the Standard Error will be.
- **Degrees of freedom (df):** The difference between the number of unique observations used as

⁸In particular, causatives do not overlap with body / mind verbs given that causatives do not specify what the causer did to bring about the result (Rappaport Hovav & Levin 2010; Lyutikova & Tatevosov 2014: 304; drawing on Shibatani 1973: 330–331), and thus cannot require a bodily / mental action.

input into the analysis ('knowns') and the number of parameters that are uniquely estimated ('unknowns').⁹

- ***t* statistic:** The coefficient (β) divided by its standard error (SE). For the intercept in Table 4.5, the *t* statistic is $(\beta=4.085) / (SE=0.0527) = 77.5$. (Note that Table 4.5 truncates the numbers 4.085 and 0.0527 to 4.09 and 0.05.)
- **Significance level (*p*):** A *p* value is the probability of finding the observed results when the null hypothesis is true. For the 'transitive' prediction in Table 4.5, $p < 0.0001$, so there would be less than a 0.01% chance of finding the observed data (where transitive verbs have strikingly lower 'each' ratings than intransitives) if there were actually *no* difference between the distributivity potential of transitives and intransitives (the 'null hypothesis'). Since *p* is so low, we can confidently reject the null hypothesis and conclude that there *is* a real difference between transitives and intransitives with respect to distributivity. Three stars (***) means $p < 0.0001$; two stars (**) means $p < 0.001$; one star (*) means $p < 0.05$; and 'n.s.' means 'not significant' (not enough evidence to reject the null hypothesis).

⁹[https://en.wikipedia.org/wiki/Degrees_of_freedom_\(statistics\)](https://en.wikipedia.org/wiki/Degrees_of_freedom_(statistics))

Table 4.5 presents this information for a model predicting a predicate’s ‘each’ rating as a function of all the independent variables in 1–5 and the interactions 6a–6c (for example, the interaction between ‘body / mind’ and ‘transitive’ is represented as ‘body / mind * trans’), allowing random intercepts for each participant and each predicate.

R command for the model reported in Table 4.5

```
lmer(each_rating ~ trans_intrans
      + bodymind
      + multilateral
      + causative
      + incremental_obj
      + trans_intrans * bodymind
      + bodymind * incremental_obj
      + causative * incremental_obj
      + (1 | SubjId)
      + (1 | full_pred),
      data = d)
```

	predicted ‘each’ rating	β	SE	df	t	p
intercept (regular intrans)	4.09	4.09	0.05	1220	77.5	***
transitive	3.51 (=4.09-0.58)	-0.58	0.05	2235	-12.3	***
body/mind	4.40 (=4.09+0.31)	+0.31	0.05	2164	6.30	***
multilateral (all intrans)	3.68 (=4.09-0.41)	-0.41	0.07	2448	-5.5	***
causative (all trans)	3.31 (=4.09-0.58-0.20)	-0.20	0.03	2185	-6.2	***
incr obj (all trans)	3.34 (=4.09-0.58-0.17)	-0.17	0.06	2113	-2.90	**
body/mind * trans	4.03 (=4.09-0.58+0.31+0.21)	+0.21	0.08	2181	2.52	*
body/mind * incr obj	3.50 (=4.09-0.58-0.17+0.16)	+0.16	0.15	2114	1.06	n.s.
causative * incr obj	3.13 (=4.09-0.58-0.20-0.17-0.01)	-0.01	0.09	2070	-0.07	n.s.

Table 4.5: Model estimates for the maximal ‘each’ model (allowing all interactions that make sense), with random intercepts for both participants and predicates.

Table 4.6 reports the same information for the ‘best’ model according to the AIC, predicting a predicate’s ‘each’ rating as a function of 1–5 and 6a (but not 6b or 6c, because the AIC comparison shows that these do not improve the model), again allowing random intercepts for each participant and each predicate.

R command for the model reported in Table 4.6

```
lmer(each_rating ~ trans_intrans
      + bodymind
      + multilateral
      + causative
      + incremental_obj
      + bodymind * trans_intrans
      + (1 | SubjId)
      + (1 | full_pred),
      data = d)
```

	predicted ‘each’ rating	β	SE	df	t	p
intercept (regular intrans)	4.08	4.08	0.05	1213	77.5	***
transitive	3.50 (=4.08-0.58)	-0.58	0.05	2236	-12.5	***
body/mind	4.39 (=4.08+0.31)	+0.31	0.05	2162	6.30	***
multilateral (all intrans)	3.67 (=4.08-0.41)	-0.41	0.07	2448	-5.5	***
causative (all trans)	3.30 (=4.08-0.58-0.20)	-0.20	0.03	2174	-6.5	***
incr obj (all trans)	3.35 (=4.08-0.58-0.15)	-0.15	0.04	2091	-3.51	***
body/mind * trans	4.05 (=4.08-0.58+0.31+0.24)	+0.24	0.08	2164	3.09	**
body/mind * incr obj	(not included)					
causative * incr obj	(not included)					

Table 4.6: Model estimates for the most parsimonious and predictive ‘each’ model according to the Akaike Information Criterion — allowing only one interaction, with random intercepts for both participants and predicates. The statistics reported below come from this model.

In parallel to Table 4.5, Table 4.7 reports the estimates, β coefficients, standard errors, degrees of freedom, t values, and significance levels (p) for a model predicting a predicate’s ‘together’ rating as a function of all the independent variables in 1–5 and 6a–6c, allowing random intercepts for each participant and each predicate.

R command for the model reported in Table 4.7

```
lmer(together_rating ~ trans_intrans
      + bodymind
      + multilateral
      + causative
      + incremental_obj
      + bodymind * trans_intrans
      + bodymind * incremental_obj
      + causative * incremental_obj
      + (1 | SubjId)
      + (1 | full_pred),
      data = d)
```

	predicted ‘together’ rating	β	SE	df	t	p
intercept (regular intrans)	3.20	3.20	0.05	1608	54.95	***
transitive	3.64 (=3.20+0.44)	+0.44	0.05	2258	8.23	***
body/mind	2.64 (=3.20-0.56)	-0.56	0.05	2196	-10.14	***
multilateral (all intrans)	3.54 (=3.20+0.34)	+0.34	0.08	2453	4.07	***
causative (all trans)	3.81 (=3.20+0.44+0.17)	+0.17	0.03	2207	4.58	***
incr obj (all trans)	3.84 (=3.20+0.44+0.22)	+0.22	0.06	2140	3.38	***
body/mind * trans	3.03 (=3.20+0.44-0.56-0.05)	-0.05	0.09	2208	-0.58	n.s.
body/mind * incr obj	2.76 (=3.20+0.44-0.56-0.32)	-0.32	0.17	2148	-1.89	n.s.
causative * incr obj	3.97 (=3.20+0.44+0.17+0.22-0.06)	-0.06	0.11	2101	-0.57	n.s.

Table 4.7: Model estimates for the maximal ‘together’ model (allowing all interactions that make sense), with random intercepts for both participants and predicates.

Table 4.8 reports the same information for the ‘best’ model according to the AIC, predicting a predicate’s ‘together’ rating as a function of 1–5 (but no interactions, because the AIC comparison shows that they do not improve the model), again allowing random intercepts for each participant and each predicate.

R command for the model reported in Table 4.8

```
lmer(together_rating ~ trans_intrans
      + bodymind
      + multilateral
      + causative
      + incremental_obj
      + (1| SubjId)
      + (1| full_pred),
      data = d)
```

	predicted ‘together’ rating	β	SE	df	t	p
intercept (regular intrans)	3.23	3.23	0.05	1271	65.54	***
transitive	3.64 (=3.23+0.41)	+0.41	0.04	2223	9.58	***
body/mind	2.62 (=3.23-0.61)	-0.61	0.04	2192	-14.34	***
multilateral (all intrans)	3.54 (=3.23+0.31)	+0.31	0.08	2452	3.86	***
causative (all trans)	3.81 (=3.23+0.41+0.17)	+0.17	0.03	2194	4.99	***
incr obj (all trans)	3.81 (=3.23+0.41+0.17)	+0.17	0.05	2117	3.38	***
body/mind * trans	(not included)					
body/mind * incr obj	(not included)					
causative * incr obj	(not included)					

Table 4.8: Model estimates for the most parsimonious and predictive ‘together’ model according to the Akaike Information Criterion — with random intercepts for both participants and predicates, but no interactions. The statistics reported below come from this model.

In sum, all the statistics reported below are drawn from the models in Table 4.6 and Table 4.8 (the ‘best’ models for the ‘each’ and ‘together’ questions, according to the AIC), isolating the effect of each independent variable. In what follows, I motivate each of these independent variables and discuss its effect on distributivity.

4.3.2 Transitive / intransitive asymmetry

Motivating the hypothesis Nearly forty years ago, [Link 1983](#) hinted at a hypothesis about a relation between argument structure and distributivity:

- (13) TRANSITIVE / INTRANSITIVE HYPOTHESIS: **Most intransitives are distributive; many verb phrases built from transitives can go both ways.**

Predicates built from most intransitive verbs (*smile*) can only be distributive, whereas predicates built from many transitive verbs (*open the window*) can be understood nondistributively.

After observing that *carry the piano* (built from a transitive verb) can be understood both distributively and nondistributively, Link writes: ‘Common nouns and intransitive verbs like *die*, however, seem to admit only atoms in their extension. I call such predicates *distributive*’ ([Link 1983](#): 132). He reiterates ([Link 1983](#): 141): ‘Most of the basic count nouns like *child* are taken as distributive, similarly IV [intransitive verb] phrases like *die* or *see*’.

Of course, we have already seen exceptions to this hypothesized pattern: *see the photo* is built from a transitive verb and is only understood distributively; *meet* is intransitive and only makes sense nondistributively; *lie* is intransitive and can be understood in both ways. But as a tendency, Link’s hypothesized transitive / intransitive asymmetry sounds plausible. To use introspective evidence, all the intransitive verbs in (14) behave like *smile* in that if Alice and Bob do these actions, then they each do (distributive).

- (14) run, swim, walk, die, blush, faint, mediate, pray, wink, laugh, arrive, disappear . . .
 ✓distributive, ✗nondistributive

In contrast, all of the predicates built from transitive verbs in (15) behave like *open the window* in that if Alice and Bob do these actions, they may do so jointly rather than individually (nondistributive). (The predicates in (15) can *also* be understood distributively, with or without covariation

depending on whether the action described by the verb can be repeated on the same object; but the important point is that they can be understood nondistributively.)

- (15) eat a pizza, write a book, send a letter, score a point, create a controversy ...
 (✓distributive), ✓nondistributive

Unlike the other hypotheses proposed below, the Transitive / Intransitive Hypothesis (13) is just a hunch, with no deep theoretical motivation. If it is indeed manifested, then we face a deeper question of why it would be so. Before taking on that question, let us test whether the Transitive / Intransitive Hypothesis is manifested empirically in the Distributivity Ratings Dataset.

Testing the hypothesis According to the combined models described above (§4.3.1; Figure 4.3), an intransitive verb is predicted to have an ‘each’ rating of 4.08, while a predicate built from a transitive verb is predicted to have a rating of 3.50 — a large difference (0.58 points on a 5-point scale), and a highly significant one ($p < 0.0001$). Turning to the model predicting the response to the ‘together’ question, an intransitive verb is predicted to have a rating of 3.23, while a predicate built from a transitive verb is predicted to have a rating of 3.64 — again, a sizable difference (0.41 points on a 5-point scale), and a highly significant one ($p < 0.0001$). As hypothesized, predicates built from transitive verbs are less distributive, and more likely to allow a nondistributive understanding, compared to intransitives.

While these findings are striking, it is much less clear how they could be explained. If the distributivity potential of a predicate is shaped by world knowledge about the event it describes, as I have claimed, then why would it also be related to whether the predicate involves an intransitive verb or a transitive one?

Perhaps it is because predicates built from intransitive verbs and transitive verbs describe different sorts of events, about which we have different world knowledge. In particular, there is converging evidence from the acquisition literature (Naigles 1990, Gropen et al. 1991, Naigles & Kako 1993), the typology literature (Dixon 1979, Hopper & Thompson 1980), and the lexical semantics

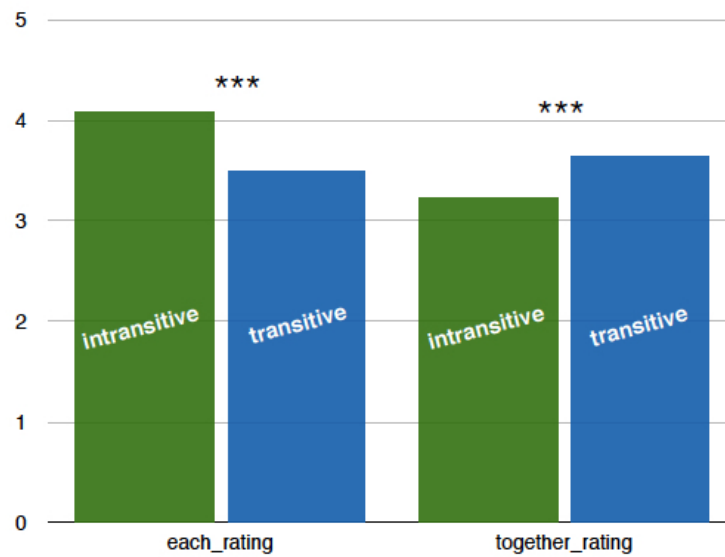


Figure 4.3: Verb phrases built from transitive verbs have systematically lower ‘each’ ratings, and systematically higher ‘together’ ratings, compared to intransitives.

literature (Dowty 1991, Croft 1991, DeLancey 1991, Levin & Rappaport Hovav 2005, Croft 2012) that transitive verbs prototypically describe events in which an agent affects another entity in some way, while intransitive verbs describe events involving only one basic participant which either acts autonomously or is affected by another entity that goes unmentioned.

In other words, the idea is that verbs with similar argument structures describe classes of events sharing certain commonalities (Fillmore 1970, Hopper & Thompson 1980, Levin 1993, Levin & Rappaport Hovav 2005). Assuming that a predicate’s potential for distributivity depends on world knowledge about the event it describes, we expect predicates describing similar sorts of events to pattern together in their potential for distributivity. Thus, I suggest that the apparent connection between argument structure and distributivity is an indirect one, driven by the types of events that tend to be described by transitive verbs versus intransitive ones.¹⁰

¹⁰See Glass 2017 for discussion; although the empirical portion of that paper is superseded by the Distributivity Ratings Dataset.

The rest of the hypotheses that I lay out aim to identify more fine-grained aspects of predicates that shape their distributivity potential. Many of these hypothesis by their nature apply disproportionately to transitives or to intransitives, indirectly helping to explain the observed asymmetry.

4.3.3 Body / mind predicates

Motivating the hypothesis *Smile* is distributive because it describes a facial action which people can only carry out individually. The same reasoning should extend to other predicates describing the actions of an individual body or mind.¹¹ Generalizing the analysis of *smile* leads to a hypothesis:

- (16) *Individuals have their own bodies and minds; so if multiple individuals carry out an action that involves one's body / mind, then they each carry out that action. Therefore:*

BODY / MIND HYPOTHESIS — Predicates describing actions that involve one body / mind are in general *only* understood distributively.

Among predicates describing bodily or mental actions, some involve intransitive verbs, while others involve transitives. Intransitive verbs describing bodily or mental actions include, for example, *smile, walk, run, sleep, faint, die, blush, blink, breathe, shrug, yawn, sneeze, sit,* and *stand* among the body verbs; and *worry, dream, fret, fume, dither,* and *cringe* among the mental / emotional ones. This category of ‘mental / emotional’ verbs includes verbs of thinking, feeling, and perceiving, but not verbs of communication or social (inter)action, such as *argue, debate, converse, date,* and *chat*, which I consider social rather than mental / emotional.

Some intransitive body / mind verbs describe events that arguably require more than one participant — perhaps one person cannot *waltz* or *tango* alone. But the majority of such verbs describe events that involve a single individual; for example, they combine easily with singular subjects (*Alice smiled / slept / worried*) in the absence of any (explicit or inferred) *with* phrase.

Such verbs of course *can* be understood distributively when applied to a plural subject. They

¹¹The world knowledge that people have their own bodies also surfaces in the literature on inalienable possession (where various languages distinguish between inherent, ‘inalienable’ possessions such as *her arm* versus transient, ‘alienable’ possessions such as *her hat*); see, for example, Guéron 2006 and references therein.

describe actions that can be carried out by one individual, so when they are predicated of multiple individuals, it is clearly possible that each individual carried out the predicate (distributive) — each individual *smiled, slept, worried*, and so on. More strongly, my claim is that these verbs not only *can* be understood as distributive, but that they largely *have* to be. (Some exceptions are discussed below.) Individuals normally have their own bodies and minds. Therefore, when a predicate describes the action of an individual body or mind, it generally *has* to distribute to each individual body / mind represented in the plural subject.

For example, all of the intransitive body / mind verbs in (17) are only understood distributively:

- (17) Alice and Bob smiled.
 slept.
 walked.
 breathed.
 fainted.
 sighed.
 blushed.
 worried.
 dreamed.
 mourned.
 meditated.

✓**Distributive:** They each did so.

✗**Nondistributive:** They did so jointly but not individually.

As for the exceptions, we have already seen one: the unusual use of *smile* from Chapter 3, repeated below. This example violates the normal assumption that the individual members of the subject each have their own body and mind; two lips do not each have their own body. Although this is a counterexample to the idea that body / mind verbs are distributive, the fact that it requires such unusual circumstances is actually compatible with the larger claim. The Body / Mind Hypothesis

(16) assumes that each member of the subject has its own body / mind; so it is not surprising that the hypothesis no longer applies when that assumption is subverted.

- (18) Alice's lips smiled (but her eyes didn't). adapted Winter & Scha 2015: 5; = (6)
- a. (??) **Distributive:** Alice's lips *each* smiled.
 - b. ✓**Nondistributive:** Alice's lips smiled jointly.

In addition to the intransitive body / mind verbs, there are also transitive ones: for example, among the body verbs, there are those involving an individual mouth: *swallow (a pill)*, *lick (a spoon)*, *bite (an apple)*. There are others requiring individual hands or feet: *grasp (a bar)*, *pinch (a child)*, *punch (a cop)*, *kick (a ball)*. Similarly, there are mental / emotional verbs describing an individual's mental processes: *see (a photo)*, *admire (a view)*, *respect (an elder)*, *smell (a rose)*, *mourn (a loss)*, *witness (a death)*, and *taste (a wine)*.

Just as among the intransitive body / mind verbs (*smile*, *sneeze*), my claim is that verb phrases built from these verbs are understood distributively: if two people *swallow a pill*, *kick a ball*, *see a photo*, or *admire a view*, then they generally each do so (distributive):

- (19) Alice and Bob swallowed a pill.

licked a spoon.

grasped a bar.

kicked a ball.

saw a photo.

admired a view.

respected an elder.

smelled a rose.

✓**Distributive:** each did so

✗**Nondistributive:** did so jointly but not individually

The prediction is that predicates involving an individual body or mind should *only* be understood distributively when applied to a plural subject.

Testing the hypothesis To test the Body / Mind Hypothesis, the first step was to tag all of the verb phrases in the Distributivity Ratings Dataset that describe bodily or mental actions. These include:

1. Verbs describing bodily actions:

- Levin’s ‘verbs of assuming a position’ (*kneel, bow, perch, slump, slouch . . .*); ‘verbs involving the body’ (*squirm, sway, twitch, wiggle, faint, breathe, sweat, vomit, weep, kneel, curtsy, snore, swallow, hiccup, sniff, sob, sleep, wink, shrug . . .*); and ‘verbs of grooming and bodily care’ (*shower, exercise, shave . . .*).
- Levin’s ‘run’ verbs (*canter, bounce, glide, hop, hurry, jog, run . . .*) and ‘modes of being involving motion’ (*tremble, waver, teeter, writhe . . .*).
- Given that individuals have their own mouths / vocal tracts, I include Levin’s ‘verbs of ingesting’ (*brunch, dine, graze, nosh, snack, swig, swallow . . .*); ‘animal sound’ verbs (*baa, bark, bay, bellow, bleat, cluck, coo . . .*); some vocal ‘performance verbs’ (*sing, intone, hum*); vocal ‘sound emission’ verbs (*scream, screech, stutter, warble*); and contact verbs requiring specific body parts (*lick, bite, punch*).

2. Verbs of emotion and perception: Levin’s ‘psych’ verbs with experiencer subjects (where the subject of the sentence is the one experiencing the relevant emotion; *admire, abhor, disdain, dislike, enjoy, envy . . .*); and ‘verbs of perception’ (*recognize, glimpse, spy, spot, view . . .*).

In total, 491 unique predicates in the Distributivity Ratings Dataset were tagged as body / mind verbs (376 intransitive, 115 transitive; see Table 4.5).

According to the models described above (§4.3.1; Figure 4.4), a regular intransitive is predicted to have an ‘each’ rating of 4.08, while a body / mind intransitive is predicted at 4.39 — a sizable effect (0.31 points) in the predicted direction, and a highly significant one ($p < 0.0001$). The

interaction between body / mind and transitivity was also significant ($p < 0.001$); a regular transitive is predicted to have an ‘each’ rating of 3.50, while a body / mind transitive is predicted at 4.05 — 0.23 points higher than if the effects of ‘body / mind’ and ‘transitive’ were kept separate.

As for the ‘together’ model, a regular intransitive is predicted to have a ‘together’ rating of 3.20, while a body / mind intransitive is predicted at 2.64 ($p < 0.0001$). This time, the interaction between body / mind and transitivity was not significant; but (just based on the main effects of transitivity and body / mind) a regular transitive is predicted to have a ‘together’ rating of 3.65, while a body / mind transitive is predicted at 3.03 ($p < 0.0001$).

In sum, body / mind predicates are more distributive and less nondistributive compared to others, as predicted by the Body / Mind Hypothesis.

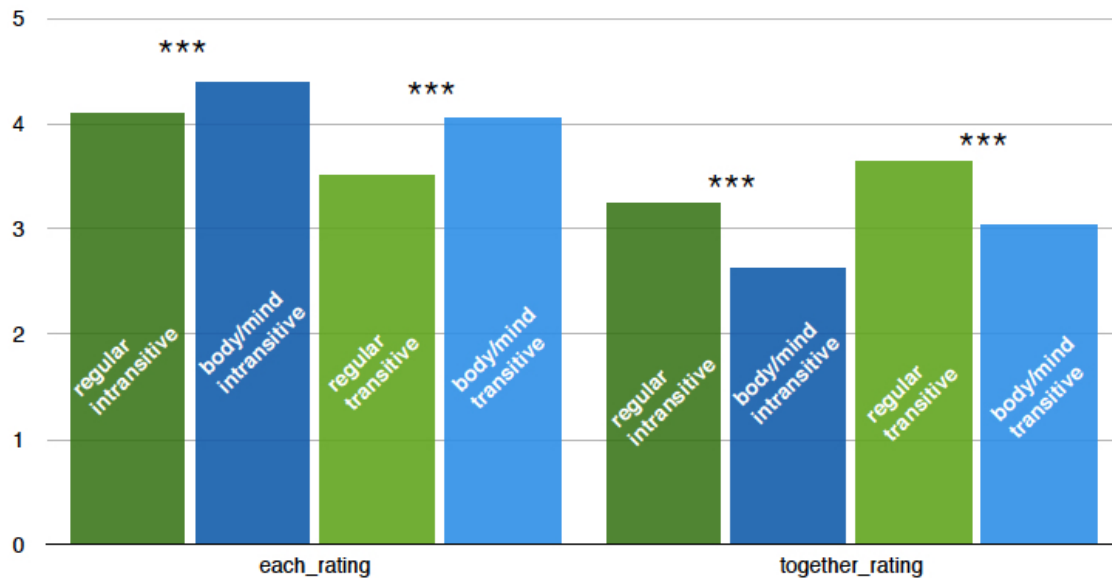


Figure 4.4: Body / mind intransitives have systematically higher ‘each’ ratings, and systematically lower ‘together’ ratings, than other intransitives. In the same way, body / mind transitives have systematically higher ‘each’ ratings, and systematically lower ‘together’ ratings, than other transitives.

Moreover, the body / mind predicates overwhelmingly involve intransitive verbs: only 23% (115 of 491) of the body / mind verbs are transitive, compared with 71% (1667 of 2338) of the verbs in

Distributivity Ratings Dataset overall. As a result, the Body / Mind Hypothesis helps to drive the observed asymmetry between transitive and intransitive verbs.

In a sense, the Body / Mind hypothesis is extremely obvious; it simply generalizes the agreed-upon analysis of *smile* to a few hundred other predicates. But in doing so, it helps to explain the distributivity potential of verb phrases at a much larger scale.

4.3.4 Multilateral predicates

Motivating the hypothesis Like *smile*, the intuitive analysis of *meet* can also be expanded. *Meet* is understood nondistributively because it describes an inherently multilateral action which individuals cannot carry out alone. Generalizing, we predict:

(20) *Some events inherently require multiple participants. Therefore:*

MULTILATERAL HYPOTHESIS — **Predicates describing events requiring multiple participants are *only* understood nondistributively.**

If an action requires multiple participants, then an individual person (*Alice*) cannot carry out that action alone (Carlson 1998, Siloni 2012). Therefore, predicates describing such actions *cannot* be understood distributively, but can *only* be understood distributively. To use introspective evidence, it seems that the predicates in (21) behave in this way:

(21) Alice and Bob met.

gathered.

dispersed.

convened.

✗Distributive: They each did so.

✓Nondistributive: They did so jointly but not individually.

As an exception, such predicates actually can be understood distributively when predicated of

plural subjects whose members are themselves groups, as in Lasersohn’s example *the committees met* (where each committee meets separately; Lasersohn 1990b: 11). Another possible exception arises if *meet* is understood to involve an unspecified implicit object (as in the somewhat unnatural dialogue: ‘who met with the invited speaker today? Alice and Bob met’.) But here again (like the *lips smiled* example (18)), the exceptions actually prove the rule. Just as lips can jointly create a smile in a way that individual humans cannot, committees but not individual humans can meet individually. Similarly, individuals can *meet* with unspecified third parties in a way that they cannot do alone. The Multilateral Hypothesis assumes that the members of the subject are atomic individuals who cannot carry out the predicate individually, so it is not surprising that the hypothesis fails when that assumption is subverted.

The Multilateral Hypothesis is arguably of a different flavor than the Body / Mind Hypothesis (§4.3.3). Body / mind predicates (*smile*) are argued to be distributive in view of the events they describe (*smiling* involves the face), combined with the fact that individuals have their own faces. In contrast, multilateral predicates are argued to be nondistributive simply in view of the events they describe (*meeting* is an inherently social / multilateral action), without relying on any further contingencies. So one might say that the distributivity of *smile* arises from a combination of lexical semantics (the nature of the event it describes) and contingent facts (that individuals have their own faces), while the nondistributivity of *meet* arises purely from lexical semantics (the nature of the event it describes). But in any case, the empirical prediction is clear: multilateral predicates should be nondistributive.

Testing the hypothesis To test the Multilateral Hypothesis, the first step is to tag all of the verbs in the Distributivity Ratings Dataset which describe inherently multilateral actions. There are some clear cases of such verbs (for example, Levin’s ‘herd’ verbs — *assemble*, *gather*, *congregate*). But there are also some fuzzier cases. Of course, one person can only *marry* or *divorce* someone else; but if two people *marry* or *divorce*, they might do so together, or perhaps might each do so with some other third party (e.g., *Alice and Bob married* when Alice married Sue and Bob married Caroline). The same goes for many other such verbs: *disagree*, *argue*, *date*, *elope*, *kiss*, *hug* ... (see Winter

2018). These verbs describe multilateral actions, but they might involve implicit third parties rather than the mutual action of the two members of the plural subject.

To make things even more uncertain, there are also verbs describing actions for which it is difficult to say whether they require multiple participants or not (Kruitwagen et al. 2017, Winter 2018). Does it really take two to *tango*, or can one *tango* (*waltz, foxtrot*) alone? What about *gossip, chitchat*, or *schmooze*? Thanks to all of these uncertainties, it is not a simple matter to code verbs for whether they describe inherently multilateral actions or not.

As an attempt to at least delineate the clearest cases, my coding for ‘multilateral’ verbs includes:

1. Levin’s ‘herd’ verbs (*group, assemble, gather, herd, convene, congregate* ...).
2. Levin’s ‘meet’ verbs (*meet, fight, battle, play* ...).
3. Levin’s ‘marry’ verbs (*marry, divorce, date, court* ...).
4. Levin’s ‘chitchat’ verbs (*chitchat, gossip, converse* ...).
5. Levin’s ‘correspond’ verbs (*war, quibble, dispute, collaborate, compete, communicate, feud, banter* ...).

In sum, 91 unique verbs were coded as describing inherently multilateral actions, all of them intransitive. (Some of the best candidates for multilateral transitive verbs, *share [a pizza]* and *coauthor [a book]*, are unfortunately not among the Levin verbs.)

Related to all of these multilateral verbs, there is also a large class of transitive, causative verbs describing events where the patient / object is required to have multiple parts: for example, Levin’s ‘mix’ verbs (*blend, combine, conjoin* ...), ‘amalgamate’ verbs (*interlock, interconnect*), and ‘disassemble’ verbs (*disconnect, unbuckle* ...). These verbs seem to describe inherently multilateral actions on the part of their objects. However, given that I have defined distributivity here only in terms of the subject of a sentence, and given that causative verbs such as *blend* were tested only in their causative form (e.g., *blend a color* as opposed to the inchoative form, *the colors blended*),

these verbs do not qualify as describing inherently multilateral actions for the purpose of the current study. For example, a person can *blend a color* individually.

According to the models described above (§4.3.1; Figure 4.5), a regular intransitive verb is predicted to have an ‘each’ rating of 4.08, while a multilateral intransitive is predicted at 3.67 ($p < 0.0001$). A regular intransitive is predicted to have a ‘together’ rating of 3.23, while a multilateral intransitive is predicted at 3.54 ($p < 0.0001$). In other words, multilateral verbs are less distributive and more nondistributive compared to other verbs, consistent with the Multilateral Hypothesis.

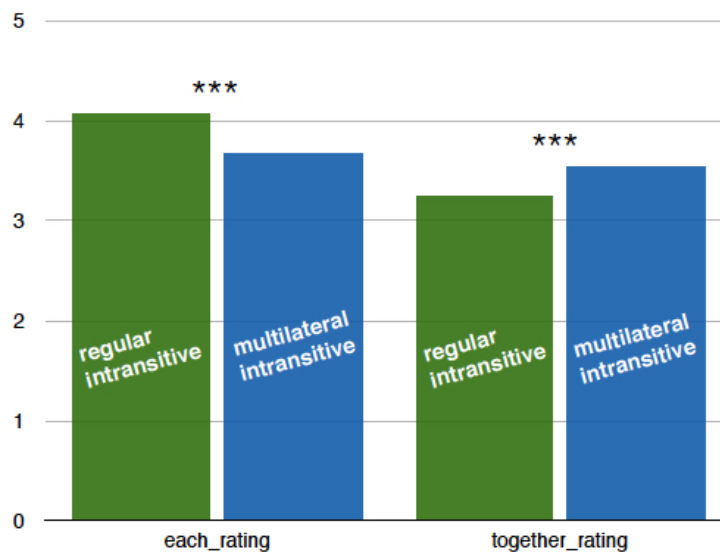


Figure 4.5: Multilateral verbs (all intransitive) have lower ‘each’ ratings and higher ‘together’ ratings than other intransitives.

In contrast to the Body / Mind Hypothesis, the Multilateral Hypothesis runs counter to the observed transitive / intransitive asymmetry. Multilateral verbs (all intransitive) are predicted to be understood *nondistributively*, in conflict with the observation that many intransitives are understood distributively. The 91 ‘multilateral’ intransitives (of 671 intransitives total) are exceptions to the generalization that intransitive verbs tend to describe events that individuals carry out individually.

4.3.5 Causatives

Motivating the hypothesis Having identified predicates that behave like *smile* in being understood distributively (§4.3.3), and like *meet* in being understood nondistributively (§4.3.4), the next goal is to identify further predicates that behave like *open the window / open a window* in being understood in both ways. While *smile* is clearly distributive because it involves the body, and *meet* is clearly nondistributive because it involves multiple parties, it is much less obvious why *open the / a window* behaves the way it does, or which other predicates should pattern with it.

My proposal is that *open* is a causative verb (Smith 1970, Dowty 1979), describing an event in which the subject causes the object to change in openness. By definition, causatives describe events of causation. I argue that this truism predicts the distributivity potential of such predicates: as a general fact about causation, it is possible for multiple individuals' actions to jointly bring about a result without each individually doing so. That, I argue, is why (22) can be understood nondistributively (22b), so that only the joint contributions of Alice and Bob together suffice to cause the opening of a window (for example, in a situation where Alice unlocks the window and Bob pushes it open).

- (22) Alice and Bob opened a window.
- a. (✓**Distributive:** They each opened a window.)
 - b. ✓**Nondistributive:** They opened a window jointly without each individually doing so.
(Alice unlocks it, Bob pushes it open.)

Generalizing, other causative predicates are predicted to behave like *open the / a window* in being able to be understood nondistributively:

- (23) *As a general fact about causation, multiple individuals' actions may be jointly sufficient but individually insufficient to cause a result. Therefore:*

CAUSATIVE HYPOTHESIS — **Causatives can be understood nondistributively**

Predicates built from transitive causative verbs (*open, lift, break*) can be understood nondistributively (as well as, perhaps, distributively — depending on definiteness and repeatability [§1.3.3]).

More formally, this hypothesis can be derived from a leading analysis of causative verbs (as foreshadowed by Dowty 1987). Causative verbs are often said to comprise a primitive building block of meaning known as CAUS, meant to express that they describe events of causation (McCawley 1968, Dowty 1979). Most influentially (in a tradition dating back to the philosopher David Hume 1748 and revived by Lewis 1973), CAUS can be defined counterfactually: the idea that an event a causes (CAUS) an event b only if b would not have happened but for a .¹²

Analyzing counterfactuals in terms of possible worlds, the counterfactual analysis states that in all of the worlds most similar to the actual world in which a does not happen, b does not happen either. In other words, if Alice opened the window, then in the closest worlds in which Alice doesn't do anything to the window, the window does not open. The counterfactual analysis has its critics (see Copley & Wolff 2014 for a review), but it makes interesting predictions about the distributivity potential of causatives (Dowty 1987).

If two events $a \wedge b$ cause a third event c , then, according to the counterfactual analysis, in the closest worlds in which $a \wedge b$ does not happen, c does not happen either. Some of the closest $\neg(A \wedge B)$ worlds might be $A \wedge (\neg B)$ worlds, or $(\neg A) \wedge B$ worlds — all predicted by the counterfactual analysis to be $\neg C$ worlds (Dowty 1987). In other words, the counterfactual analysis of causation captures an intuition: that two factors may be jointly sufficient, but individually insufficient, to cause a result.

On such an analysis, a sentence such as (24) means that Alice and Bob did something which caused the window to open.

(24) Alice and Bob opened the window.

¹²A technical note: Lewis defines causation as a relationship between events, but uses propositions rather than events in order to pick out the correct worlds for his counterfactual analysis. For him, an event a causes an event b if all of the closest not- A worlds are not- B words — where A is defined as the proposition that the event a occurs, and B is the proposition that the event b occurs. I follow Lewis in using lower-case letters for events and capital letters for the propositions that those events occur.

The event of Alice and Bob doing something can be decomposed into an event of Alice doing something, and another event of Bob doing something. In the closest worlds in which nothing is done to the window by Alice or Bob, the window does not open. Some of these worlds may be ones in which Alice or Bob does something to the window alone, but the window still does not open in these worlds. In other words, the individual contributions of Alice and of Bob may be separately insufficient, but jointly sufficient, to cause the window to open — giving rise to a nondistributive understanding of the predicate.

Theoretically, this logic should extend to all causative predicates, predicting all of them to allow a nondistributive understanding (in addition to whatever distributive understandings are available depending on the definiteness of the object and the repeatability of the action; §1.3.3). To use introspective data, I argue that this is the intuition behind the nondistributive understanding of all of the causative predicates in (25): that Alice and Bob somehow realized the result upon the object through their combined efforts.

- (25) Alice and Bob opened the window.
 lifted the table.
 collapsed the tent.
 moved the statue.
 removed the stain.
 angered the committee.
 debunked the rumor.
 beautified the room.
 melted the chocolate.
 doubled the revenue.
 shortened the skirt. . . .

(✓**Distributive:** They each did so.)

✓**Nondistributive:** They jointly caused the result without each individually doing so.

As further introspective motivation for this hypothesis, we can explore its predictions. In some cases, the same verb can be understood as either causative or non-causative (Levin & Rapaport Hovav 2014): *clean* can be understood as ‘causing something to become clean’, or as ‘carrying out some prototypical actions associated with cleaning’, such as vacuuming or dusting, without entailing that its object becomes clean. We therefore predict that when a predicate built from *clean* is understood as causative, it must allow a nondistributive understanding; but when it is *not* understood as causative, it might only make sense distributively.

This prediction is indeed consistent with the data: the causative (26) can be understood nondistributively, in a situation in which Alice and Bob only jointly make the stove clean — for example, if Alice sprays it with degreaser and Bob wipes it off. In contrast, it is much more difficult to imagine a nondistributive understanding of the non-causative (27): if Alice and Bob did some apartment-cleaning (dusting, vacuuming, and so on), we normally infer that they each did so.

- (26) **Causative:** Alice and Bob cleaned the stove (so that it was spotless when they finished).
- a. ✓**Distributive:** each cleaned it (perhaps on different occasions).
 - b. ✓**Nondistributive:** cleaned it jointly without each individually doing so.
- (27) **Non-causative:** Alice and Bob cleaned the apartment (for awhile; but it was still messy when they stopped).
- a. ✓**Distributive:** each did some apartment-cleaning.
 - b. ?? **Nondistributive:** jointly did some apartment-cleaning without each doing so.

The contrast between (26) and (27) illustrates that a predicate’s distributivity potential does not just depend on the specific verb involved, but is further shaped by whether that verb is construed as causative, consistent with the Causative Hypothesis.

Furthermore, the hypothesis should not just apply to lexical causatives such as *open the window*, but is also predicted to extend to periphrastic causatives such as those in (28), on the assumption that these also describe events in which Alice and Bob cause a change upon the object. And indeed,

periphrastic causatives seem to allow a nondistributive understanding, just as lexical causatives do.

- (28) a. Alice and Bob caused the window to open.
 b. Alice and Bob got the window open.
 c. Alice and Bob made the window open.
 (✓**Distributive**: each caused it to open — perhaps on different occasions).
 ✓**Nondistributive**: caused it to open only jointly.

To sum up, causative predicates describe a unified class of events — those in which the subject causes a change upon the object. As a general fact about causation (predicted by the counterfactual analysis), the actions of multiple agents may be individually insufficient, but jointly sufficient, to cause some result. We therefore hypothesize that predicates built from causatives should allow a nondistributive understanding. This Causative Hypothesis is supported by preliminary evidence from introspective data. We predict that it should be manifested quantitatively in the Distributivity Ratings Dataset as well.

Testing the hypothesis in the Distributivity Ratings Dataset To test the Causative Hypothesis, the first step was to tag the verbs in the dataset as ‘causative’ or non-‘causative’. Of course, only transitive verbs can be considered causative in the relevant sense of causing a change to be realized on the object. While there is no agreed-upon list of all the causative verbs, it seems clear that any verb undergoing the ‘causative / inchoative alternation’ (*break the vase / the vase broke*) should count as causative — encompassing for example Levin’s long list of change-of-state verbs (*break, shatter, increase, boil*). Even non-alternating verbs can be considered causative if they entail that their object underwent a change of state: for example, the ‘remove’ verbs (which entail that their object is removed in some way: *purge, void, confiscate*); similarly the ‘put’ verbs (which entail that their object ends up in a certain location — *pocket, jail* — or that something else is put on or in the object: *pollute, soak, shroud*), and the ‘psych’ verbs describing events where the subject causes the object to feel some emotion (*annoy, frighten*; Belletti & Rizzi 1988). In total, 945 of the 1667

transitive verbs in the dataset were coded as causative.

According to the models described above (§4.3.1; Figure 4.6), a regular (non-causative) transitive is predicted to have an ‘each’ rating of 3.50, whereas a causative transitive is predicted at 3.30 ($p < 0.0001$). A regular (non-causative) transitive is predicted to have a ‘together’ rating of 3.64, whereas a causative transitive is predicted at 3.81 ($p < 0.0001$). In other words, causatives are less distributive and more nondistributive than other transitives, consistent with the Causative Hypothesis.

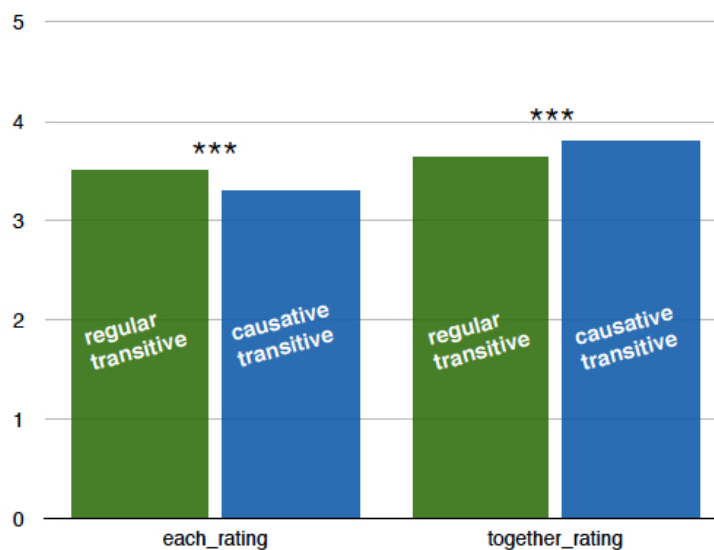


Figure 4.6: Causatives (all transitive) have lower ‘each’ ratings and higher ‘together’ ratings than other transitives.

With 945 (57%) of the 1667 transitive verbs in the dataset labeled as causative, this finding constitutes a far-reaching pattern in the distributivity potential of verb phrases. Moreover, because causatives as defined here inherently involve transitive verbs, the fact that causatives can be understood nondistributively helps to explain the observed tendency for predicates built from transitive verbs to allow a nondistributive understanding (§4.3.2).

Of course, the Distributivity Ratings Dataset faces the limitation that every transitive verb is

tested with a particular object, which may itself contribute to the inferences drawn about the predicate's distributivity potential. For example perhaps *clean a house* differs from *wipe a skillet* not just because *clean* is causative and *wipe* is not, but also because houses are larger than skillets, so that cleaning a house might require more participants than wiping a skillet. However, in the aggregate, the difference between *clean a house* and *wipe a skillet* should not matter. The procedure for choosing objects (§4.2.1) is not expected to give causatives and non-causatives systematically different sorts of objects in a way that would bias their distributivity potential. Moreover, by treating each predicate as a random effect, the regression models that I conducted control for arbitrary differences between individual verb-object combinations. However *clean a house* differs from *wipe a skillet* in particular, the statistical analysis finds a robust difference between causatives and non-causatives in general, consistent with the Causative Hypothesis.

4.3.6 Predicates with incremental objects

Motivating the hypothesis Alongside the hypothesis that causatives can be understood nondistributively, I also hypothesize that incremental-object predicates (Tenny 1987, Krifka 1989, Dowty 1991, introduced in Chapter 2) can be understood nondistributively as well:

- (29) *As a result of the incremental homomorphism between the parts of an object and the parts of the event (eat the pizza), multiple individuals might each carry out the verb event on a different portion of an incremental object (might eat a different portion of the pizza), only adding up to the whole (eating the whole pizza) between them. Therefore:*

INCREMENTAL HYPOTHESIS — **Incremental-object predicates can be nondistributive.** Predicates with objects construed as incremental (*eat the pizza*, where the full pizza is consumed at the end of the event) can be understood nondistributively (as well as, perhaps, distributively — depending on definiteness and repeatability [§1.3.3]).

As explained above (§2.4), incremental-object predicates describe events in which the parts of the object correspond to the parts of the event: in (30), there is a homomorphism between the pizza

and the event of eating it, so that when the pizza is half gone, the event of eating it is half over, and when the pizza is gone, the event of eating it is over.

(30) Alice ate the pizza.

Incremental-object predicates are not always distinguished from causatives; while Dowty 1979 predates the concept of incremental-object predicates, he discusses many such predicates (*paint a picture*) under the guise of ‘accomplishments’ in the sense of Vendler 1967 — and suggests that all accomplishments are causative. Conversely, Rothstein 2004 subsumes many causative predicates (*repair the computer*) under the category of accomplishments, which she analyzes as inherently incremental. Such analyses blend causatives and incremental-object predicates together. But there are compelling arguments for distinguishing these two classes. Causative predicates entail a result (*break the vase* entails that the vase is broken), while incremental-object predicates need not: *read the book* does not entail any change in the book (Rappaport Hovav 2008). Causative verbs usually cannot be used with implicit objects (*I broke* cannot be used to convey *I broke stuff*), while incremental-object verbs often can (*I ate* is roughly equivalent to *I ate stuff*; Rappaport Hovav 2008). Incremental-object predicates are atelic with mass objects (*eat some cake* is atelic, because the unboundedness of *cake* does not place an endpoint on the event of eating it; Verkuyl 1972, Krifka 1989), while causatives can be understood as telic even with mass objects (*break some glass* can be telic; Levin 2000).

Some predicates can be placed into both classes, such as *cube the zucchini* (from the Distributivity Ratings Dataset): the zucchini is causally affected, in that it is cut into cubes, but may also be understood to be incrementally affected, in that each portion of the zucchini may correspond to a different part of the event of cubing it.

(31) The chef cubed the zucchini.

But there are also predicates that only fit into one class or the other: *read the book* has an incremental object but is not causative; *calm the baby* is causative but does not have an incremental

object (the baby is not calmed piece by piece). Thus, these two classes are treated as overlapping, but distinct.

By definition, incremental-object predicates describe events in which the parts of an object correspond to parts of the event described by the predicate. I argue that this fact can be used to predict their distributivity potential. Informally, it is always possible for multiple individuals to each carry out the event described by the verb on a different portion of the object, only jointly adding up to the whole. As a result, incremental-object predicates with subjects denoting multiple individuals should allow a nondistributive understanding. For example, (32) can be understood nondistributively, as in (32b), so that only between the two of them did Alice and Bob fully consume the pizza.

- (32) Alice and Bob ate a pizza.
- a. **(✓Distributive):** Each ate a [different] pizza.)
 - b. **✓Nondistributive:** Ate one pizza between them.
(Alice eats one half of the pizza, Bob eats the other half.)

In contrast to (32), predicates without incremental objects such as (33) might *only* be understood distributively, with no available nondistributive understanding:

- (33) Alice and Bob saw a photo.
- a. **✓Distributive:** They each saw a photo.
 - b. **✗Nondistributive:** They saw a photo jointly without each individually doing so.

More formally, this hypothesis can be derived from the assumption (Chapter 2) that verbs and thematic roles are cumulative. Recall from Chapter 2 that a verb such as *eat* can be analyzed as a set of eating events, as in (34) (where events are represented as tuples consisting of a label for the event and its thematic roles). Then, for any two events e_1 and e_2 in this set, the cumulativity assumption requires that their sum $e_1 \oplus e_2$ is also in this set. The sum of two *eat* events is also an *eat* event; its

agent is the sum of the agent of e_1 and the agent of e_2 , and its theme is the sum of the theme of e_1 and the theme of e_2 . Again, this setup guarantees the natural result that if Alice eats half the pizza and Bob eats the other half, then Alice and Bob eat the full pizza between them.

$$(34) \quad \llbracket eat \rrbracket = \{ \langle e_1, agent = Alice, theme = half\ the\ pizza_1 \rangle, \\ \langle e_2, agent = Bob, theme = half\ the\ pizza_2 \rangle, \\ \langle e_1 \oplus e_2, agent = Alice \oplus Bob, thm = half\ the\ pizza_1 \oplus half\ the\ pizza_2 \rangle \}$$

Whenever a predicate's object is construed as incremental in this way, we predict the predicate to allow an understanding in which the members of the subject each carry out the event described by the verb on a different portion of the object. If the extension of an incremental-object verb includes an event of Alice and Bob carrying out the event described by the verb on the full object, it is always possible for the extension of the verb to also include a subevent of Alice carrying out the verb event on one part of the object and Bob carrying out the verb event on the rest, as in (34): each eating different portions of the pizza, adding up to the whole between them.

We therefore predict that incremental-object predicates should be able to be understood nondistributively (29). Using introspective data, I argue that this intuition explains the nondistributive understandings of the predicates in (35): that Alice and Bob each performed the action described by the verb on a different portion of the object, with their contributions only adding up to the whole between them.

- (35) Alice and Bob wrote the book.
 ate the pizza.
 painted the wall.
 ran the marathon.
 copy-edited the document.
 built the Lego castle.
 searched the house.

vacuumed the basement.

loaded the truck.

recited the poem. . . .

(✓**Distributive**: They each did so.)

✓**Nondistributive**: They did so jointly but not individually (by each doing a different part).

As further motivation for the Incremental Hypothesis, we consider cases in which the same predicate may or may not be understood as telic (Dowty 1979, Krifka 1992, Rothstein 2001, Smollett 2005, Rappaport Hovav 2008). With *read the magazine*, perhaps the magazine is fully read over the course of the event (telic) — or perhaps only some arbitrary portion of the magazine is read (atelic). The analysis predicts that when the predicate is construed as telic, it should allow a nondistributive understanding (because each member of the subject carries out a part of the event on a different part of the object, jointly adding up to the whole); whereas when the predicate is construed as atelic, it might *only* have a distributive understanding (because if two people read some arbitrary portion of a magazine, then they each also read some arbitrary portion of that magazine).

And indeed, the telic incremental-object predicate in (36) can be understood nondistributively, for example if Alice reads one half of the magazine and Bob reads the other. In contrast, it is much more difficult to imagine a nondistributive understanding of (37), in which the magazine is not fully read (atelic). Given that people have their own mental processes, then if Alice and Bob did some magazine-reading, we generally infer that they each did.

(36) **Telic**: Alice and Bob read the magazine (from start to finish, to check it for errors).

a. ✓**Distributive**: They each read it.

b. ✓**Nondistributive**: They each read part of it, only jointly reading the whole thing.

(37) **Atelic**: Alice and Bob read the magazine (for awhile, but didn't finish it).

a. ✓**Distributive**: They each did some magazine-reading.

b. (??) **Nondistributive**: They jointly did some magazine-reading, without each indi-

vidually doing so.

Consistent with the Incremental Hypothesis, the contrast between (36) and (37) shows that what matters most, even more than the specific predicate involved, is the construal of its object.

This hypothesis is predicted to extend to all cases in which a verb's object is construed as incremental. Sometimes, we ascribe incrementality even to the objects of verbs that are not typically classified as incremental-object verbs — particularly when the object is a numeral plural (Krifka 1992). For example, *see* is not a prototypical incremental-object verb (the subparts of a *see-the-zebra* event do not correspond to subparts of *the zebra*); but an event in which Alice sees seven zebras can be split into subevents in which each individual zebra is seen, culminating when seven zebras are seen in all (Krifka 1992).

Normally, *see* — even though it is a transitive verb — only allows a distributive understanding with a definite singular object: if Alice and Bob *see the zebra*, we generally infer that they each do so. But when its object can be construed as incremental, as in *see seven zebras*, we predict a nondistributive understanding to be systematically available.

As predicted, (38) can be understood nondistributively — for example, in a situation in which Alice sees three zebras and Bob sees four more (Krifka 1992). Again, the incremental construal of the object is more important for a predicate's distributivity than the particular verb involved.

- (38) Alice and Bob saw seven zebras. adapted Krifka 1992: 43
- a. ✓**Distributive:** each saw seven zebras.
 - b. ✓**Nondistributive:** saw seven zebras between them.

Summing up, incremental-object predicates describe a unified class of events — those in which there is a homomorphism between the parts of the object and the parts of the event described by the predicate. As a general fact about such events (predicted by the assumption that verbs and thematic roles are cumulative), multiple agents may each individually carry out the verb event on a different subpart of the object, only jointly adding up to the whole. Based on this theoretical discussion and

introspective examples, we predict that (29) should be manifested quantitatively in the Distributivity Ratings Dataset.

Testing the hypothesis in the Distributivity Ratings Dataset To test the hypothesis that predicates with incremental objects can be understood nondistributively while those without incremental objects may only be understood distributively, the first step was to tag predicates for whether their objects can be construed incrementally or not. (Of course, only transitive verbs can have incremental objects.)

In contrast to the causatives, it is full verb phrases, not just individual verbs, which can be construed incrementally (for example, *eat a pizza* can be construed incrementally, while *eat pizza* cannot; Krifka 1989). Also, most verbs are either causative or not¹³, while verb phrases such as *eat a pizza* might be construed as telic (if the pizza is fully consumed by the end of the event), or might be construed as atelic (if only some arbitrary portion of the pizza is eaten — see Krifka 1992, Jackendoff 1996, Rothstein 2001, Smollett 2005, Rappaport Hovav 2008); and the Incremental Hypothesis only applies to the telic construal (see (36)–(37)). Moreover, predicates built from the same verb might or might not be construed as incremental depending on the size of the object (Rappaport Hovav 2008): when someone *eats a grape*, the grape may be eaten all at once, so that its parts do not correspond to the parts of the eating event (non-incremental), even though other predicates built from *eat* do involve an incremental mapping between the object and the event (*eat a pizza*). For these reasons, coding the ‘incremental’ predicates is a rather subtle matter.

There is no agreed-upon list of incremental-object predicates, so I had to construct one myself. The main categories of incremental predicates include:

1. (Physical or intellectual) consumption predicates: for example, those built from Levin’s ‘verbs of ingesting’, such as *devour a fish*, *ingest a drug*, *guzzle a beer*, and *consume a fish*; or, more metaphorically, ‘learn’ verbs such as *read an article* and *memorize a poem*.

¹³An exception to the idea that verbs can be clearly classified as causative or not: the causative and non-causative construals of *clean* discussed in §4.3.5.

2. Creation predicates: for example, those built from ‘image-creation’ verbs (*etch a glass, illustrate a book, write a book*); ‘coloring’ verbs (*glaze a biscuit, lacquer a box*); and ‘build’ verbs such as *build a house, assemble a sandwich, and carve a statue*.
3. Spatial-coverage predicates: for example, *iron a shirt, weed a garden, inspect a facility, seed a field*.

In sum, 201 (12%) of 1667 predicates built from transitive verbs were coded as incremental.

According to the models described above (§4.3.1; Figure 4.7), a regular (non-incremental) transitive is predicted to have an ‘each’ rating of 3.50, whereas a transitive with an incremental object is predicted at 3.35 ($p < 0.0001$). A regular (non-incremental) transitive is predicted to have a ‘together’ rating of 3.64, whereas a transitive with an incremental object is predicted at 3.81 ($p < 0.0001$). In other words, incremental-object transitives are less distributive and more nondistributive than other transitives, consistent with the Incremental Hypothesis. (See the Appendix 6.3 for a followup experiment offering further evidence consistent with this hypothesis.)

To summarize, it was hypothesized that the structure of a telic incremental-object event allows that multiple individuals may each carry out the verb event on a different portion of the object, only adding up to the whole between them (giving rise to a nondistributive understanding of the verb phrase). This hypothesis is manifested in the Distributivity Ratings Dataset. With 201 of the 1667 transitive verbs in the dataset labeled as having potentially incremental objects, this finding constitutes a substantial pattern in the distributivity potential of verb phrases. Moreover, because incremental-object predicates inherently involve transitive verbs, their behavior helps explain why predicates built from transitive verbs are more likely to allow a nondistributive understanding than intransitives (§4.3.2).

4.3.7 Discussion

The analysis of *smile, meet, and open the window* has been generalized to predict the distributivity potential of a large number of verb phrases, and these predictions have been found to be manifested

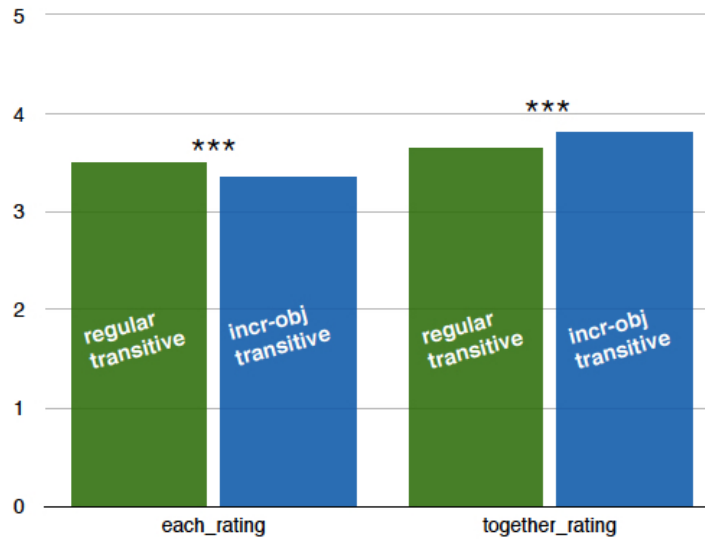


Figure 4.7: Predicates with objects that can be construed as incremental (all built from transitive verbs) have lower ‘each’ ratings and higher ‘together’ ratings than other transitives.

empirically.

In a sense, it is hardly shocking that other body / mind predicates behave like *smile*, or that other multilateral predicates behave like *meet*. But we began with three predicates (*smile*, *meet*, *open the window*) and now predict the distributivity of 1637 predicates (476 body / mind predicates, 91 multilateral predicates, 945 causatives, and 125 incremental-object predicates that are neither body/mind nor causative). These 1637 predicates constitute 70% of the total 2338 tested: substantial progress.

There is, of course, more work to be done. For example, the Body / Mind Hypothesis predicts all predicates describing the actions of individual bodies and minds to be understood distributively; but there are further non-body / mind predicates that also behave that way. Spatial location predicates (*arrive*, *depart*, *exit / enter the room*) do not require an individual body or mind; but in general, if two individuals are located at a particular place, then they are *each* located at that place (subparts share the location of the whole: if Bill is in Texas, then Bill’s brain is in Texas; [Schwarzschild 1996](#):

Chapter 5). Therefore, such spatial predicates are predicted to be distributive: if two people *arrive* or *enter a room*, then they each do so. So although the Body / Mind Hypothesis covers several hundred predicates, there are others that it leaves out.

Similarly, there are further predicates which behave like causatives and incremental-object predicates in being understood nondistributively as well as distributively. *Rent* is neither causative nor incremental (renting something does not cause that thing to change, nor does it incrementally affect that thing), and yet if two people *rent a car*, perhaps they each do so (distributive), or perhaps they do so jointly (nondistributive) — presumably because renting involves possession, and individuals can possess things individually or jointly (an explanation which extends to *buy*, *own*, *sell*, *lease*, and so on). Thus, while causatives and incremental-object predicates constitute large and diverse classes of predicates that can systematically be understood nondistributively, they are not the only ones to do so.

There are more patterns to be found. But this chapter charts a path for studying the distributivity potential of verb phrases in a systematic manner.

4.4 Chapter summary

This chapter has put forward a series of far-reaching hypotheses about the distributivity potential of various types of verb phrases, which are theoretically motivated based on independent facts about the types of events described by these predicates. These hypotheses are empirically supported in a large new dataset.

This study constitutes the literature’s first attempt to systematize the distributivity potential of verb phrases at a large scale. Backed by quantitative evidence, the question of ‘which predicates are understood in which ways and why?’ becomes a realm of concrete investigation. The cover analysis from Chapter 3 leaves a predicate’s distributivity potential to ‘what we know about the event’ it describes; this chapter has taken on the task of explaining what aspects of ‘what we know about the event’ matter and why.

Chapter 5

Adjectives

Having identified aspects of events that shape the distributivity potential of the verb phrases describing them (Chapter 4), this chapter takes up the same goal among adjectives.¹ As in the realm of verb phrases, different adjectives are understood in different ways with respect to distributivity, but it is an open question which ones are understood in which ways and why. On the assumption that a gradable adjective relates an individual to its measurement (‘degree’) along a scale (Bartsch & Vennemann 1972, Seuren 1973, Cresswell 1976, Rullmann 1995, Kennedy 1999), I argue that the understandings available to a gradable adjective are predicted by the measurement-theoretic properties of the scale it invokes (Stevens 1946, Suppes & Zinnes 1962, Krantz et al. 1971, Krifka 1989, Schwarzschild 2002, Schwarzschild 2006, Sassoon 2007, Sassoon 2010, Lassiter 2011, Solt 2015, Lassiter 2017): how the measurement of the composite $a \oplus b$ relates to the measurements of its constituent parts a and b individually.

5.1 Introduction

In the literature and in this dissertation, most discussion of distributivity has involved verb phrases — *smile, meet, open the window*. But the same phenomenon arises among adjectives in predicative

¹A version of this chapter is published as [Glass 2018](#).

position (Schwarzschild 1996, Schwarzschild 2006). Some are understood only distributively (1), others only nondistributively (2) (at least, if we don't reinterpret *connected* to mean *connected to* some sort of implicit object); still others can be understood in both ways (3).

- (1) The boxes are new.
- a. ✓**Distributive:** Each box is new.
 - b. ✗**Nondistributive:** The boxes are jointly new but not individually so.
- (2) The boxes are connected.
- a. ✗**Distributive:** Each box is connected.
 - b. ✓**Nondistributive:** The boxes are jointly connected but not individually so.
- (3) The boxes are heavy. adapted Schwarzschild 1996
- a. ✓**Distributive:** Each box is heavy.
 - b. ✓**Nondistributive:** The boxes are jointly heavy but not individually so.

In addition to these three categories (1)–(3) which are familiar from the discussion of verb phrases, there is also a fourth category among adjectives: those that could plausibly be understood nondistributively, but which in reality strongly prefer to be distributive (Quine 1960, Schwarzschild 2011). We can imagine a nondistributive understanding for (4) — i.e., that the combined height of a stack of boxes qualifies as *tall* although each individual box is short or of average height. But it is much more natural for (4) to convey that *each* box is tall (distributive). Schwarzschild 2011 names these predicates ‘stubbornly distributive’, on the grounds that they ‘stubbornly’ refuse to be understood nondistributively, even though they theoretically could be. (Table 5.1 lays out this typology, leaving out the *connected* type, which I set aside.)

- (4) The boxes are tall. adapted Schwarzschild 2011: 3
- a. ✓**Distributive:** Each box is tall.
 - b. **Nondistributive (imaginable, but not easily available):** The boxes are jointly tall but

not individually so.

Distributive	✓ Dist. : Each new	<i>The boxes are new.</i>	✗ Nondist. : Jointly new
Both ways	✓ Dist. : Each heavy	<i>The boxes are heavy.</i>	✓ Nondist. : Jointly heavy
‘Stubbornly distributive’	✓ Dist. : Each tall	<i>The boxes are tall.</i>	?? Nondist. : Jointly tall

Table 5.1: Distributivity potential of different types of adjectives.

As among verb phrases, it largely remains an open question which adjectives behave like *new*, like *connected*, like *heavy*, or — adding the ‘stubbornly distributive’ ones to the mix — like *tall*. Of course, as among verb phrases, presumably an adjective’s distributivity potential is somehow grounded in what we know about the property it describes: *new* describes age; boxes have their own ages, so if two boxes are *new*, they each are. Presumably *connected* is nondistributive because it involves a sense of reciprocity not shared by the other adjectives (which is why I do not discuss the *connected* type further).

In a recent advance, [Scontras & Goodman 2017](#) have claimed that *heavy* (which can be understood in both ways) differs from *tall* (stubbornly distributive) because the joint height of boxes depends on the transitory way they are arranged, while the joint weight of boxes is stable (§5.2) — proposing a pragmatic explanation for what might otherwise appear to be a lexical idiosyncrasy, just as I aim to do here. But a more fundamental question remains open: for which adjectives is a nondistributive understanding imaginable, whether it is available or not? *Tall* could theoretically be understood nondistributively (4b), even though this understanding is not easily available. For *heavy*, both understandings are imaginable and available (3). Whatever the difference between *tall* and *heavy*, there is also a question of what separates these two predicates from *new*, for which it is difficult to even imagine a nondistributive understanding (1b). What separates the adjectives above the double line in Table 5.1 from those below it?

To capture the difference between *tall* and *heavy* on the one hand, and *new* on the other hand, this chapter proposes an account using measurement theory (§5.3). The idea is that for a gradable

adjective A to have a nondistributive understanding, the measurement along the scale encoded by A of two things together $\mu(a \oplus b)$ must be able to exceed the measurement of each thing individually ($\mu(a)$ and $\mu(b)$). Then the contextual standard for what counts as A can be set in such a way that $a \oplus b$ exceeds the standard for A while a and b individually fall short of it — a nondistributive understanding, because the adjective A is true of $a \oplus b$ together, but not of a or b alone. Depending on the behavior of the particular scale associated with the adjective, this ordering might or might not be possible, explaining which adjectives can or cannot be understood nondistributively.

As with the verb phrases (Chapter 4), the strategy is to identify the features of reality that shape the distributivity potential of a predicate describing it. The difference here is that while the distributivity potential of a verb phrase depends on features of the event it describes, the distributivity potential of a gradable adjective depends on properties of the scale it invokes. Different types of predicates derive their distributivity potential in different ways; but it is never arbitrary.

5.2 Literature on the distributivity of adjectives

Schwarzschild gives *large*, *round*, *big*, and *long* as examples of ‘stubbornly distributive’ adjectives. He analogizes them to count nouns such as *cat*, in that both stubbornly distributive predicates and count nouns apply only to individuals, not pluralities; but as for why these adjectives in particular behave as stubbornly distributive, he leaves that as a ‘mystery’ (Schwarzschild 2011: 5).

5.2.1 A pragmatic explanation for *heavy* versus *tall*

To explain why *heavy* can be easily understood nondistributively while *tall* ‘stubbornly’ prefers to be distributive, Scontras & Goodman 2017 observe that the joint weight of boxes is stable, while their joint height depends on the transitory way they are arranged (in a stack versus side by side).

They describe an experiment in which a robot named Cubert is responsible for handling boxes at a factory. The boxes either come out of the box-dispensing machine in a regular stack, or in a haphazard manner (the ‘random’ condition). Each time, Cubert describes the boxes to his friend

Dot, saying *The boxes were heavy / tall / big*, and so on. Experimental participants were asked whether Cubert intended to describe the boxes as a whole ('collective'), or individually ('distributive'). Scontras and Goodman find that *tall* and *big* are more likely to be understood nondistributively when the boxes come out of the dispenser in a predictable manner than when they come out haphazardly, while *heavy* is not affected by the arrangement of the boxes.

Instead of stipulating that *tall* and *big* are 'stubbornly distributive' while *heavy* is 'complaisantly' nondistributive, Scontras and Goodman derive this distinction pragmatically: hearers will not expect a speaker to intend *tall* to be nondistributive, given that the joint height of boxes is transitory and unstable; while hearers may expect a speaker to intend *heavy* to be nondistributive, since the joint weight of boxes is consistent. As predicted by this analysis, when the joint height of boxes is more stable (when the boxes come out of the dispenser in a regular stack), the nondistributive understanding of *tall* accordingly becomes more available.² *Heavy* is not influenced by the arrangement because the joint weight of boxes does not depend on it.³

In a further experiment, Scontras & Goodman 2017 test 25 different dimensional adjectives (5), grouped by the dimension that they measure (depth, height, and so on) along with the direction in which they measure it (increasing versus decreasing). For example, *tall* can be said to measure height in an increasing direction: taller things have *more* height (Seuren 1978, Kennedy 2001). *Short* measures height in a decreasing direction: shorter things have *less* height.

²While it is possible for *tall* to be understood nondistributively with enough context (e.g., in a situation where boxes regularly come out of a machine in a stack), there are other 'stubbornly distributive' adjectives where the imaginable nondistributive understanding is much more elusive. For example, even with a context favoring a nondistributive understanding, Syrett 2015 finds experimentally that *the boxes are round* is robustly rejected to describe square boxes arranged into a round circle (presumably for the reason that Scontras and Goodman propose: the joint shape of boxes depends on their transitory spatial arrangement while their individual shape does not). But Scontras and Goodman's analysis is still consistent with the finding that *tall* is more pragmatically pliant than *round*. They do not predict that every 'stubbornly distributive' adjective will become nondistributive with enough context, as *tall* does; instead, they predict that nondistributive understandings are more available for adjectives describing properties of groups that are stable with respect to arrangement.

³Another insight from their experiment: Cubert also either moves all the boxes together on a dolly ('move'), or inspects them ('inspect'). Participants are less willing to choose the distributive understanding of *heavy* in the 'move' condition, which Scontras and Goodman say is because participants infer that Cubert does not know how much each individual box weighs when he moves them all together, while he might know if he inspects them. Given that speakers should only make claims for which they have evidence (Grice 1989), the idea is that experimental participants consider Cubert's evidence when trying to figure out what he means — a different type of pragmatic effect on distributivity.

(5) **Dimensional adjectives studied by Scontras & Goodman**

<i>tall</i>	height	increasing
<i>short</i>	height	decreasing
<i>deep</i>	depth	increasing
<i>flat</i>	height	decreasing
<i>low</i>	height	decreasing
<i>big</i>	size	increasing
<i>small</i>	size	decreasing
...

Scontras and Goodman find that for size and height adjectives, the increasing-direction ones (*big, tall*) are more likely to be understood nondistributively (‘collectively’) than the decreasing-direction ones (*small, short*), particularly in the condition where the boxes come out of the dispenser in a regular stack. In other words, the nondistributive understanding (6b) is more easily available than (7b).

(6) The boxes were {big / tall }.

- a. **Distributive:** Each box is { big / tall }.
- b. **Nondistributive:** The boxes together are { big / tall }, but not individually.

(7) The boxes were { small / short }.

- a. **Distributive:** Each box is { small / short }.
- b. **Nondistributive:** The boxes together are { small / short }, but not individually.

For Scontras & Goodman 2017: 304, this contrast arises because (6b) is more likely to be true than (7b):

‘It seems unlikely that Cubert would intend to communicate that a stack of boxes taller than him is collectively short when the distributive alternative is available, namely that each box is short ... When an interpretation appears unlikely to be true (e.g., describ-

ing a tall stack of boxes as collectively short), listeners are unlikely to attribute that interpretation to speakers' utterances.'

As a result, Scontras and Goodman say, *small* and *short* behave as if they are 'stubbornly distributive' — not just because the collective height or size of boxes is unstable, but also because a stack of boxes is unlikely to be considered *short* or *small*. (This explanation is not entirely convincing, though; gradable adjectives such as *small* and *short* are famously vague, so it is surprising that a stack of boxes could not be considered *short* or *small* compared to what Cubert expected, even if the stack is taller / larger than Cubert himself.)

In sum, Scontras and Goodman provide a convincing pragmatic explanation for Schwarzschild's observation that the nondistributive understanding of certain 'stubbornly distributive' adjectives like *tall* is imaginable but not easily available. But many questions remain open.

5.2.2 Open questions

In order for an adjective such as *tall* to be 'stubbornly distributive', it must have an imaginable-but-pragmatically-unavailable nondistributive understanding. For *heavy* to be 'complaisantly' nondistributive, it must also have an imaginable (and pragmatically available) nondistributive understanding. It is still an open question which adjectives have such an understanding. For other adjectives such as *new*, a nondistributive understanding is very difficult to even imagine. So what separates *new* (only distributive) from *tall* and *heavy* (for which a nondistributive understanding is imaginable, whether or not it is available)?

On the one hand, there is evidence that the distributivity potential of adjectives is systematically related to the nature of the properties they describe. Just as we can imagine a nondistributive understanding (pragmatically available or not) for *heavy* and *tall*, the same goes for other adjectives that describe physical dimensions in an increasing direction (*large*, *big*, *wide*, *long*). The fact that semantically similar adjectives pattern together suggests that their behavior is tied to their meaning.

On the other hand, the distributivity potential of some adjectives appears idiosyncratic. Many adjectives come in antonym pairs such as *heavy* / *light*, *open* / *closed*, and *tall* / *short*, describing

inversely related properties (Sapir 1944, Cruse 1976, Cresswell 1976, Seuren 1978, Lehrer & Lehrer 1982, Muehleisen 1997, Kennedy 2001). Some antonym pairs pattern together with respect to distributivity: *open* and *closed* are both distributive only (8), in that if multiple boxes are open or closed, they each are. *Beautiful* and *ugly* can both be understood in both ways (9): multiple boxes might each be beautiful or ugly (distributive), or might only be so when arranged together (nondistributive).

- (8) The boxes are {open / closed}.
- a. ✓**Distributive:** Each box is {open / closed}.
 - b. ✗**Nondistributive:** The boxes together are {open / closed}, but not individually.
- (9) The boxes are {beautiful / ugly}.
- a. ✓**Distributive:** Each box is {beautiful / ugly}.
 - b. ✓**Nondistributive:** The boxes together are {beautiful / ugly}, but not individually.

Presumably the two halves of these antonym pairs pattern together because of the nature of the properties they describe: separate containers can only be *open* or *closed* individually; aesthetic judgments can be made about individual objects or collections thereof. But then it is surprising that there are also antonym pairs which diverge in their potential for distributivity. While *heavy* can be understood in both ways (3), it is quite difficult to imagine how its antonym *light* could be true of multiple boxes without also being true of each one.

- (10) The boxes are light.
- a. ✓**Distributive:** Each box is light.
 - b. ✗**Nondistributive:** The boxes together are light, but not individually.

On the surface, it is puzzling that some antonym pairs pattern together while others diverge. Within this apparent idiosyncrasy, however, there is again the hint of a pattern: the antonym pairs that diverge tend to be dimensional adjectives (*heavy / light*, *tall / short*, *big / small*), and it is always

the decreasing-direction one that prefers to be distributive — again suggesting that this behavior can be somehow tied to the similarities between these adjectives.

Scontras and Goodman describe adjectives such as *light* (decreasing-direction adjectives that are understood distributively) as ‘stubbornly distributive’, like *tall* — suggesting that *light*-type adjectives have an imaginable-but-pragmatically-unavailable nondistributive understanding. But while it is clear how two boxes could be considered *short* individually and *tall* when stacked, it is not at all clear what it would mean for two boxes to be *heavy* individually and *light* together. Rather than grouping *light* with *tall* as Scontras and Goodman do, I would argue that *light* behaves like *new* in that it is difficult to even imagine how it could be understood nondistributively: if two boxes together are *light*, then they each are (distributive).

Summarizing again, it is an open question which adjectives behave like *new*, *connected*, *heavy*, or *tall*, and why. Similar adjectives (*tall*, *large*, *big*) behave similarly, suggesting that it is not totally arbitrary; but it is not clear what creates these patterns, nor why some antonym pairs pattern together while others diverge. Scontras and Goodman’s convincing proposal for why *tall* differs from *heavy* is only part of the story. One would also want to know which adjectives are like *tall* and *heavy* in having an imaginable nondistributive understanding, and which are like *new* and *light* in only making sense distributively. Parallel to the investigation of verb phrases in Chapter 4, the goal of the chapter is to explain how the distributivity potential of these adjectives is derived from the properties they describe.

5.3 Background on gradable adjectives and measurement theory

I focus on gradable adjectives (Bartsch & Vennemann 1972, Seuren 1973, Cresswell 1976, Klein 1980, Kennedy 1999, Kennedy 2007) — adjectives that can be degree-modified (*very tall*, *somewhat heavy*) and participate in comparative constructions (*more beautiful*, *less full*).

Gradable adjectives such as *heavy* are commonly analyzed as measure functions, mapping individuals to degrees along a scale measuring the property described by the adjective (Cresswell 1976, von Stechow 1984, Rullmann 1995, Kennedy 1999 *et seq*). *Heavy* applied to *the box* is analyzed to

return its degree of weight — for example, 25lbs.

$$(11) \quad \text{heavy}(\text{the box}) = 25\text{lbs}$$

Ultimately (11) should yield a truth value, so when *heavy* is used as a predicate in its basic form (as opposed to as a comparative or superlative), we need some additional material to map 25lbs into something that can be true or false. The idea is that (11) is true if the box's weight exceeds some contextual standard θ for what counts as *heavy* in the context (Cresswell 1976), which in turn depends on what the box is being implicitly compared to (Klein 1980): if 25lbs exceeds the standard θ_{heavy} .

$$(12) \quad \llbracket \text{The box is heavy} \rrbracket = 1 \text{ iff } \text{heavy}(\text{the box}) \geq \theta_{\text{heavy}}$$

(As for the difference between ‘relative’ gradable adjectives such as *heavy* and ‘absolute’ gradable adjectives such as *clean*⁴ — Unger 1978, Rotstein & Winter 2004, Kennedy & McNally 2005, Kennedy 2007, Lassiter & Goodman 2013 — the idea is that both types have the same semantics, but that the contextual standard θ for a relative adjective is less certain than for an absolute adjective, because relative adjectives are associated with unbounded scales while absolute adjectives are associated with bounded ones. In addition to the measurement-theoretic properties of scales explored below, boundedness represents another way that the nature of a scale shapes the behavior of an adjective.)

On the assumption that a gradable adjective is defined in terms of a scale, I propose that the distributivity potential of adjectival predicates can be explained in terms of the structure of this

⁴For background, relative gradable adjectives are those like *heavy* and *tall*: they are interpreted relative to some comparison class — a heavy book is lighter than a heavy car. They are also vague: it is difficult to pinpoint a standard for what counts as *heavy*; there are ‘borderline cases’ where it is difficult to decide whether an object of intermediate weight should count as *heavy* or not; and such adjectives participate in the Sorites Paradox (attributed to Eubulides of Miletus; Hyde & Raffman 2014), whereby we accept that any box one gram lighter than a *heavy* box should still count as *heavy*, resulting in the absurd conclusion that a weightless box is *heavy*. In contrast, absolute gradable adjectives are those like *clean*, *empty*, *open*, and *closed*: they do not depend as heavily on a comparison class, and are less vague, seeming not to allow borderline cases and being less susceptible to the Sorites Paradox. Relative gradable adjectives such as *heavy* are associated with ‘open’ scales (there is no limit to how *heavy* something could be), while absolute gradable adjectives are associated with ‘closed’ scales (when something is totally free of dirt and germs, it can get no cleaner).

scale, which can be characterized using measurement theory.

Measurement theory (Stevens 1946, Suppes & Zinnes 1962, Krantz et al. 1971, Krifka 1989, Schwarzschild 2002, Schwarzschild 2006, Sassoon 2007, Sassoon 2010, Lassiter 2011, Solt 2015, Lassiter 2017) is a mathematical system used to analyze different sorts of measurements (height, weight, time, temperature, likelihood, and so on; see Chapter 2 of Lassiter 2011 and Lassiter 2017 for a thorough overview which inspires the discussion here). Rather than taking numbers as foundational to measurement, measurement theory begins from the qualitative notion of relative ordering (which Sapir 1944 takes as psychologically basic): for two objects a and b in a domain, does a outrank b with respect to the property P that is being measured? Does b outrank a (Lassiter 2011)?

This qualitative ranking is then mapped to the natural numbers in such a way that all and only the information from the qualitative ranking is preserved. The natural numbers are not foundational, but only derived as a way of quantitatively reflecting the original qualitative ranking.

The reason for not taking the natural numbers as basic is that the natural numbers support operations and relations that certain qualitative rankings do not support. The natural numbers are structured by their ratios to one another — one hundred is twice as large as fifty — while not all measurement systems support such a structure. If it is 100 degrees Fahrenheit in Washington, D.C. and 50 degrees Fahrenheit in Chicago, it does not strictly make sense to say that D.C. is *twice as hot* as Chicago. One reason why not: temperature could just as well be measured in degrees Celsius (Lassiter 2011), in which case it is 38 degrees Celsius in D.C. and 10 degrees Celsius in Chicago, which would mean that D.C. is 3.8 times as hot as Chicago, rather than twice as hot. Measurement theory makes it possible to construct different sorts of scales with different attributes, using only as much structure from the natural numbers as suits the property being measured.

To map qualitative rankings into the natural numbers without introducing more structure than desired, measurement theory invokes a homomorphism μ . μ relates a qualitative structure $\langle X, \succeq_P \rangle$ (where X is the domain of objects, and \succeq_P ranks one object above another with respect to the property P) to a quantitative structure $\langle \mathbb{R}, \geq \rangle$ (where \mathbb{R} is the domain of real numbers and \geq ranks one number as greater than or equal to another). For all x, y in the domain X , it is required that

(taken from Lassiter 2011 p. 33):

- $\mu(x) \in \mathbb{R}, \mu(y) \in \mathbb{R}$
(μ maps x and y into the real numbers)
- If $x \succeq_P y$, then $\mu(x) \geq \mu(y)$
(μ preserves the ordering given by \succeq_P)

The qualitative structure may also contain further operations, whose structure must also be preserved when mapped by μ into the real numbers. For the study of distributivity, the most important of these operations is the ‘concatenation’ operation \circ , which takes two objects a and b and returns a composite object $a \circ b$. (As long as there is no overlap between a and b , the concatenation operation \circ is equivalent to the join operation \oplus from Link 1983; Lassiter 2011: Chapter 2 building on Krifka 1989.) Depending on the structure of the scale, $\mu(a \circ b)$ might bear different relationships to $\mu(a)$ and $\mu(b)$.

For a scale such as weight, μ preserves the structure of the natural numbers, including the way they can be added together and their ratios to one another. The weight of Box A and Box B together (concatenated) is equivalent to the weight of Box A plus the weight of Box B; $\mu(a \circ b) = \mu(a) + \mu(b)$. Moreover, if $\mu(\text{Box A})$ is 50lbs and $\mu(\text{Box B})$ is 25lbs, then Box A is twice as heavy as Box B (a ratio which is preserved in the metric system, unlike the temperature ratios discussed above: converting pounds to kilograms, Box A weighs 22.6kg and Box B weighs 11.3kg — still twice as heavy). A scale with these properties is called an ‘additive’ scale because the addition operation $+$ can be used to handle concatenation, or a ‘ratio scale’ because it preserves ratios.

Additive scales can be subsumed under a larger class of ‘positive’ scales — those where $\mu(a \oplus b)$ is guaranteed to exceed $\mu(a)$ and $\mu(b)$. For additive scales, $\mu(a \oplus b)$ is equivalent to $\mu(a) + \mu(b)$, but other positive scales do not meet this strict definition. A 50-decibel trumpet combined with a 50-decibel piano does not amount to 100 decibels, but rather to something around 53 decibels⁵;

⁵Source: Quora post written by Berkeley physics professor Richard Muller; <https://www.quora.com/Will-two-separate-50-dB-sounds-together-constitute-a-100-dB-sound>

two sounds together are louder than each one individually, but not in an additive manner. Similarly, cost is generally positive but not necessarily additive: a \$100 shirt and \$100 pants might cost \$200 (additive), or perhaps there is a sale ('buy one, get one 50% off') so that the total cost is only \$150 (positive, but not additive).

In contrast, temperature is generally neither additive nor positive (at least when we restrict our attention to the thermometer temperature of non-chemically-reactive substances — see below for discussion of different construals of temperature); $\mu(a \circ b)$ is not equivalent to $\mu(a) + \mu(b)$. As Lassiter points out, if the soup in one bowl a is 75 degrees Fahrenheit, and the soup in another bowl b is 100 degrees Fahrenheit, then the concatenation of the two soups (poured together into a larger bowl) is certainly not 175 degrees Fahrenheit, but instead comes out to an intermediate temperature between 75 and 100 degrees (depending on the relative volumes of the soups). This type of scale is called 'intermediate' because $\mu(a \circ b)$ falls between $\mu(a)$ and $\mu(b)$.

Based on the way a and b relate to their concatenation $a \circ b$, one can define a variety of scales (13) — (among others) additive (13a), positive (13b), intermediate (13c). We can also define 'atom-only' scales (13d), which Lassiter explicitly connects to predicates that can only be distributive: it does not make sense to measure the extent to which two concatenated individuals $a \circ b$ together are *sick*, given that only individuals can be sick. While other scales are classified by how $\mu(a \oplus b)$ relates to $\mu(a)$ and $\mu(b)$, here $\mu(a \oplus b)$ is undefined.

(13) **Some types of scales** adapted / abridged Lassiter 2011: 45

a. **Additive:** $\mu(a \circ b) = \mu(a) + \mu(b)$

Example: weight

b. **Positive (of which additive is a special case):** $\mu(a \circ b) \succ \mu(a), \mu(a \circ b) \succ \mu(b)$

Example: loudness, cost

c. **Intermediate:** If $a \succeq b$, then $\mu(a) \succ \mu(a \circ b) \succ \mu(b)$

Example: temperature (of non-chemically-reacting substances, measured by a thermometer)

- d. **Atom Only:** \succeq_P contains no concatenations; i.e. $a \succeq_P b$ implies that a, b are atoms.

Example: a predicate like *sick*, which only makes sense applied to individuals

Before proceeding, it is worth noting that the superficially formal and precise concatenation operation \circ (or equivalently, assuming no overlap, \oplus) actually requires some context-dependent, entity-specific decisionmaking about how the composite object $a \oplus b$ is to be assembled from its constituent parts a and b . To measure the height of two boxes $a \oplus b$, do we measure them as a stack (in which case height is additive with respect to concatenation) or side by side and take the height of the taller one (in which case height is not additive)? To measure the temperature of two soups, do we mix them together or leave each one in its own container? If they are mixed together, could they react with one another chemically? What if the two elements being combined are of different types; what would it mean to concatenate a soup and a box? For current purposes, my approach is simply to articulate how I take concatenation to work for the different composite entities that I discuss.

The next step is to use the classification in (13) to derive the distributivity potential of adjectives from the way $\mu(a)$ and $\mu(b)$ relate to $\mu(a \oplus b)$ along the scale associated with the adjective.

5.4 Explaining the distributivity potential of adjectives

With this background, I turn to adjectives predicated of plurals, which is where distributivity comes in. Combining the semantics from Chapter 3 with the analysis of gradable adjectives, (14) requires *heavy* to be individually true of each cell of a contextually supplied cover of *the boxes* — meaning that the degree of weight of each cell in the cover of *the boxes* exceeds some contextual standard θ for what counts as *heavy* in the context.

- (14) The boxes are heavy.

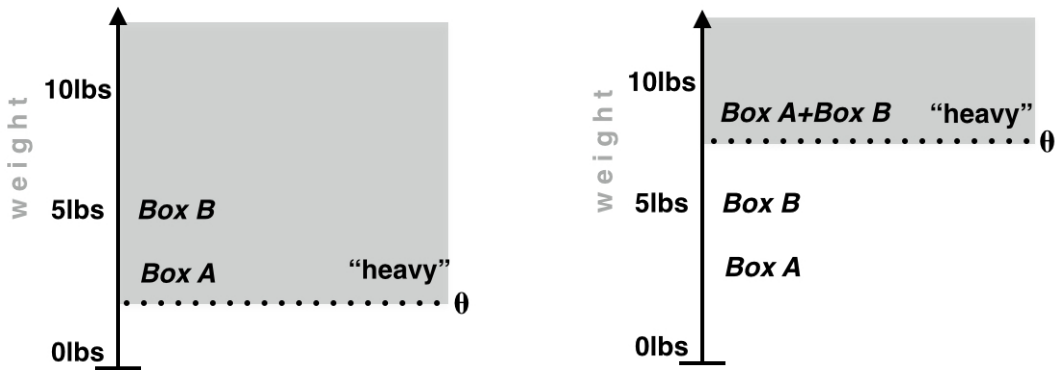


Figure 5.1: Distributive and nondistributive understandings of *heavy*.

$$\forall x[x \in Cov(\text{the boxes}) \rightarrow \text{heavy}(x) \geq \theta]$$

- a. **Distributive:** Each box is heavy
 $Cov = \{ \{a\}, \{b\} \}$
- b. **Nondistributive:** The boxes are jointly but not individually heavy
 $Cov = \{ \{a, b\} \}$

In a context with only two boxes a and b , (14) is understood distributively if each box is placed in its own cell (14a), nondistributively if both boxes occupy the same cell (14b). We have already seen that *heavy* can be understood in both of these ways, so both of these cover settings are plausible. The semantic analysis of gradable adjectives helps explain why. Imagine that Box A weighs 3lbs and Box B weighs 5lbs. For each box to individually qualify as *heavy*, each box's weight must exceed the contextual threshold θ_{heavy} for what counts as *heavy* in the context (the left side of Figure 5.1; the gray zone represents everything that is considered *heavy*, exceeding θ along the weight scale).

For the two boxes to qualify as *heavy* jointly but not individually, the weight of $a \oplus b$ must exceed θ (so that the two boxes together are considered *heavy*), while the weight of each individual box falls short of θ (so that each individual box is not considered *heavy*). Weight is additive; the weight of $a \oplus b$ is the weight of a plus the weight of b (assuming that a weighs 3lbs and b weighs 5lbs, then $a \oplus b$ weighs 8lbs). Therefore, if the contextual standard for *heavy* is set at 7lbs, then

$a \oplus b$ exceeds it while a and b each fall short of it — nondistributive (the right side of Figure 5.1).

Of course, since θ_{heavy} depends on a comparison class (Klein 1980), it could theoretically be set at different levels when weighing a single box (compared to other individual boxes) versus when weighing a pair of boxes (compared to other pairs of boxes). But when *heavy* is understood nondistributively — when two boxes qualify as *heavy* while each individual box does not — it actually seems that the individual boxes a and b and the pair of boxes $a \oplus b$ are all compared to the same consistent standard (for example, ‘what I can carry easily’). Otherwise, if Box A and Box B together are considered *heavy* relative to other pairs of boxes, then it seems likely that Box A and Box B would also each be considered *heavy* relative to other individual boxes. Such a variable setting of θ_{heavy} makes it much harder to imagine how *heavy* could be construed nondistributively.

Generalizing the discussion of *heavy*, I argue that, when θ is held constant in this way, then:

- (15) CLAIM: For a gradable adjective A to be understood nondistributively, $a \oplus b$ together must exceed a and b individually on the scale invoked by A . That way, the standard θ for what counts as A can be set so that $a \oplus b$ exceeds θ while a and b individually fall short of it.

Depending on the measurement-theoretic properties of the scale associated with the adjective, this ordering may or may not be possible, shaping the distributivity potential of that adjective. In particular, only adjectives associated with ‘positive’ scales — those where $\mu(a \oplus b)$ exceeds $\mu(a)$ and $\mu(b)$ — can fulfill (15).

We predict that adjectives describing similar sorts of properties should pattern together with respect to distributivity, on the assumption that they all display the same measurement-theoretic behavior. So it becomes possible to handle large classes of adjectives all at once.

Adjectives have not been organized into classes as fine-grained as those that Levin 1993 offers for verb phrases, but they can be broadly categorized using the system of Dixon 1982 (and Dixon 2004). Aiming to explain how the grammatical category of ‘adjective’ relates typologically to the conceptual category of ‘properties’, which adjectives are thought to describe, Dixon presents seven classes of ‘Property Concepts’ that he finds are encoded as adjectives in languages with an open

adjective class.

(16)	Dixon's seven classes of Property Concepts	Dixon 1982
	dimension	<i>big, small, long, tall, short, wide, deep, ...</i>
	age	<i>new, young, old, ...</i>
	value	<i>good, bad, lovely, perfect, ...</i>
	color	<i>black, white, red, ...</i>
	physical	<i>hard, soft, heavy, wet, hot, sour, ...</i>
	speed	<i>fast, quick, slow, ...</i>
	human propensity	<i>proud, jealous, happy, kind, brave, ...</i>

This classification is not meant to be exhaustive (there are adjectives that do not fit easily into it, such as *healthy, sick, abstract, or philosophical*); it does not account for a distinction between increasing and decreasing adjectives (*tall vs. short*); and some of the adjectives' classifications are debatable (perhaps *heavy* might be considered 'dimensional' rather than 'physical', especially since it patterns with many other dimensional adjectives in measuring a property that is additive with respect to concatenation). But the Dixon system constitutes a starting point for grouping adjectives by their meaning. When we explain the distributivity potential of one adjective, the Dixon system helps to identify others which can be handled in the same way.

Increasing-direction dimensional adjectives It was pointed out above (§5.2.2) that increasing-direction dimensional adjectives such as *tall, big, and heavy* constitute the clearest exemplars of adjectives with an imaginable nondistributive understanding, whether this understanding is easily available (as for the 'complaisantly nondistributive' *heavy*) or not (as for the 'stubbornly distributive' *tall* and *big*). Measurement theory helps to explain why. Like weight, height and size are additive ($\mu(a \oplus b) = \mu(a) + \mu(b)$), so that $a \oplus b$ is guaranteed to exceed a and b along these scales (Figure 5.2). That way, just as for *heavy*, the contextual standard θ for what counts as *tall* or *big* can be set so that $a \oplus b$ surpasses it while a and b fall short of it individually. On the proposed analysis (15), that is what gives these adjectives their imaginable nondistributive understanding.

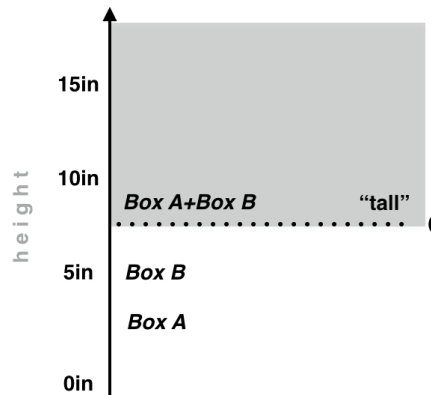


Figure 5.2: The boxes together qualify as *tall*, but not individually.

As for why *tall* and *big* tend to be understood distributively even though they have an imaginable nondistributive understanding, I echo the proposal of Scontras and Goodman (§5.2): that the joint height or size of boxes is not stable enough for the speaker and hearer to coordinate on. But now we also understand why these adjectives have an imaginable nondistributive understanding, even if it is pragmatically inaccessible.

This explanation extends not just to all of the increasing-direction dimensional adjectives, but also adjectives describing other properties associated with positive scales, such as *expensive*, *long* (in the sense of duration as well as physical length), and *loud*.

Decreasing-direction dimensional adjectives In contrast to increasing-direction dimensional adjectives such as *big*, *heavy* and *tall*, it is difficult to even imagine a nondistributive understanding for decreasing-direction adjectives such as *light*, *short*, and *small* (§5.2.2); it is not clear what it would mean for a pair of boxes to be jointly *light* while individually *heavy*. Here too, measurement theory helps to explain why. Weight is additive, so the weight of $a \oplus b$ together exceeds the weight of a and b individually. This ordering is what makes it possible for *heavy* to be understood nondistributively (Figure (14)); but the same property prevents *light* from being understood that way.

Light measures weight in a decreasing direction: a *lighter* box has *less* weight. Thus, the addi-

tivity of weight means that a and b are individually *lighter* than $a \oplus b$ together, which is the reverse of the ordering that would be needed for a nondistributive understanding (15). It is impossible to set a contextual standard θ for what counts as *light* so that $a \oplus b$ exceeds it while a and b fall short of it individually (Figure 5.3), explaining why these decreasing-direction dimensional adjectives differ from their antonyms in only being understood distributively.

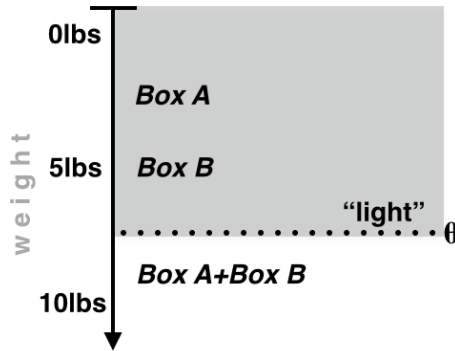


Figure 5.3: *Light* is true of things lighter than the contextual standard θ (here, 7lbs).

Other adjectives with the same behavior include *cheap*, *short* (height, duration), and *quiet*.

Adjectives with scales that are intermediate with respect to concatenation As noted above, the scale of temperature (at least, the thermometer temperature of non-chemically-reacting substances) is intermediate with respect to concatenation: $a \oplus b$ falls between a and b . Based on the claim in (15), we predict that an adjective associated with an intermediate scale should not be able to be understood nondistributively, because $a \oplus b$ does not surpass a or b individually (Figure 5.4).

As predicted, the temperature adjectives *warm* and *cold* only make sense distributively (17)–(18). Cake and fudge together are no warmer than they are individually, so they cannot jointly qualify as *warm* without also each doing so.

- (17) The cake and the fudge are warm.
- a. ✓**Distributive:** The cake is warm, the fudge is warm.

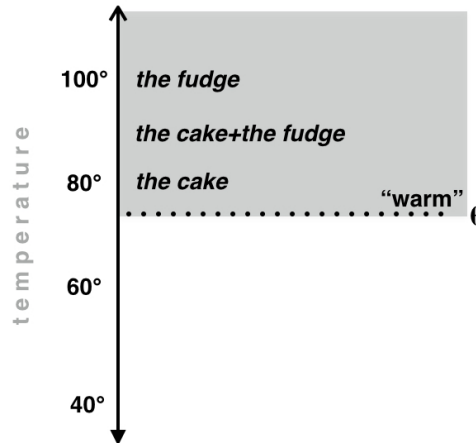


Figure 5.4: The cake and the fudge together are no *warmer* than they are individually.

- b. **✗Nondistributive:** The cake and fudge together are warm but not individually.
- (18) The cake and the ice cream are cold.
- a. **✓Distributive:** The cake is cold, the ice cream is cold.
 - b. **✗Nondistributive:** The cake and ice cream are cold together but not individually.

This explanation also extends to many of Dixon’s ‘speed’ adjectives (*fast, slow*) and ‘physical’ adjectives (*hard, soft, wet, dry*), also intermediate with respect to concatenation.

The behavior of the scale with respect to concatenation matters more than the adjective This discussion of temperature has been explicitly restricted to the thermometer temperature of non-reactive substances. Why non-reactive substances? Because if two chemicals react with one another to produce heat, then the temperature of the two chemicals together may exceed the temperature of each one. Why only thermometer temperature? Because temperature can be construed in different ways — as an objective numerical measurement; or as a subjective bodily experience, perhaps the tactile temperature of a specific object, or the ambient temperature of a room, or one’s body temperature in relation to the comfortable range for humans (Koptjevskaja-Tamm & Rakhilina 2006,

Koptjevskaja-Tamm 2011). Construed in these ways, temperature may not be intermediate with respect to concatenation: (19) may convey that a person only feels warm (comfortable in cold weather) when wearing a hat and a scarf together, not just one or the other — a nondistributive understanding of *warm*.

- (19) The scarf and the hat are warm.
- a. ✓**Distributive:** The scarf is warm, the hat is warm.
 - b. ✓**Nondistributive:** The scarf and hat are warm together but not individually.

In other words, depending on the nature of concatenation (whether it involves a chemical reaction or not) and on the construal of temperature (thermometer vs. subjective experience), the temperature scale behaves differently with respect to concatenation. In turn, the way temperature behaves with respect to concatenation influences the distributivity potential of temperature adjectives such as *warm*. When temperature is intermediate with respect to concatenation, *warm* is only understood distributively. When it is additive, *warm* can be understood nondistributively, because $a \oplus b$ can be considered *warm* while a and b individually do not qualify.

It was observed in Chapter 4 that the distributivity potential of a given verb phrase depends on whether it is construed as causative or not (*clean the apartment*), or whether it is construed as telic or not (*read the magazine*), so that specific lexical items are less important for distributivity than the construal of the events they are taken to describe. In the same way, the distributivity potential of an adjective is not a lexical fact about the specific adjective, nor is it fully predicted from the property (e.g., temperature) measured by it. Instead, its distributivity potential depends on the behavior of its scale with respect to concatenation (intermediate versus positive). As predicted by the proposed analysis, what is most important is the way $a \oplus b$ relates to a and b along this scale.

Adjectives with scales that are irregular with respect to concatenation So far, the adjectives discussed in this chapter have mostly described properties that can be measured objectively (height, weight, temperature). But other adjectives describe more subjective properties, such as predicates

of personal taste (*delicious, pretty, disgusting*; Lasersohn 2005).

To predict the distributivity potential of these adjectives, one would need to know how their associated scales behave with respect to concatenation: how does the deliciousness of $a \oplus b$ relate to the deliciousness of a and of b ? There is no single answer to this question. Chocolate is delicious and coffee is delicious, and together they are even better. Chocolate is delicious and salmon is delicious, but together they are disgusting. Because there is no rhyme or reason to what people consider *delicious* (in contrast to what count as *heavy* or *tall*), there is no pattern to the way these predicates behave with respect to concatenation.

As a result, all subjective predicates are predicted to allow a nondistributive understanding, because it is possible for $a \oplus b$ to exceed a and b individually along the scale associated with the adjective. (It is also possible for $a \oplus b$ to fall below a and b ; subjective predicates are so irregular that anything can happen). This prediction seems correct; the distributive understandings of (20)–(21) are certainly more natural, but the nondistributive understandings can be imagined as well:

(20) The flowers are {pretty, ugly}.

- a. ✓**Distributive:** Each flower is {pretty, ugly}.
- b. ✓**Nondistributive:** The flowers together are {pretty, ugly}, but not individually.

(21) The appetizers are {delicious, disgusting}.

- a. ✓**Distributive:** Each appetizer is {delicious, disgusting}.
- b. ✓**Nondistributive:** The appetizers together are {delicious, disgusting}, but not individually.

The rest of Dixon's subjective 'value' adjectives (*good, bad, perfect*) behave in the same way.

'Atom-only' adjectives Most of the adjectives discussed so far in this chapter have described properties that can be instantiated by pluralities as well as by individuals. Two boxes together have height, weight, temperature, and beauty. But, as Lassiter notes when he lays out the different

types of scales (13), there are also adjectives describing properties that can only be instantiated by individuals, such as *sick*. Like the body / mind verbs discussed in Chapter 4 (*smile, die, blush*), being *sick* involves the body; individuals have their own bodies, so they can generally only be *sick* individually. Similarly, given that individuals have their own mental processes, they can only be *depressed, worried, or religious* as individuals.

The proposed analysis (15) predicts that an adjective can be understood nondistributively if $a \oplus b$ together exceeds a and b individually along the scale invoked by the adjective. But perhaps it is bizarre to even measure the *sickness* of two people together. In that case, adjectives like *sick* can only be distributive, not just because $\mu(a \oplus b)$ does not exceed $\mu(a)$ or $\mu(b)$, but because $\mu(a \oplus b)$ is not defined in the first place.

This analysis extends not just to all bodily adjectives such as *sick, dead, awake, and alive*, but also to many of Dixon's 'human propensity' adjectives that describe emotions (*proud, jealous*), on the assumption that emotions are experienced individually. It further encompasses certain adjectives describing location (*local, close, nearby*) and origin (*American*), based on the spatial fact that if two individuals are located somewhere or are from somewhere, then they each are (§4.3.7).

Hard-to-classify adjectives Finally, there are many adjectives for which it is difficult to assess their distributivity potential, as well as the measurement-theoretic properties argued to shape it. It was observed above (§5.1) that *new* only makes sense distributively. But is that because newness is an 'atom-only' property like *sick*, which can only apply to individuals (because entities have their own ages)? Or is it a property that is 'intermediate' with respect to concatenation, so that the *newness* of two things together falls in between the *newness* of each one? The result is the same either way — *new* is distributive — but the reason is not entirely clear.

More generally, as we move away from adjectives measuring objective properties such as height and width, judgments become fuzzy.⁶ The more elusive the scale evoked by the property, and the more uncertain its behavior with respect to concatenation, the more indeterminate its distributivity potential seems to be, which is in fact also consistent with the proposed analysis.

⁶Perhaps even the behavior of *new* is flexible: can old boxes be considered jointly *new* if arranged in a new way?

Discussion This chapter has offered an explanation for the distributivity potential of gradable adjectives. The data that has been covered is not as extensive as the verb phrases discussed in Chapter 4, but the measurement-theoretic analysis makes quite general predictions.

We now understand why increasing-direction dimensional adjectives such as *heavy* make the best examples of adjectives that can be understood nondistributively (and why *tall* could imaginably be nondistributive, even if it prefers not to be): because these adjectives invoke scales that are additive with respect to concatenation, meaning that $a \oplus b$ is guaranteed exceed a and b individually. We also understand why some antonym pairs behave differently from one another (*heavy* vs. *light*), while others pattern together (*new / old*, *clean / dirty*; *pretty / ugly*), all captured by the way these different scales behave with respect to concatenation (Table 5.2).

Distributive	<i>Box A & Box B are new (light, short, full, empty).</i> ✓Dist.: Each new ✗Nondist.: Jointly new because $a \oplus b$ can't exceed a, b on <i>new</i> scale
Both ways	<i>Box A & Box B are heavy (expensive, beautiful, ugly).</i> ✓Dist.: Each heavy ✓Nondist.: Jointly heavy because $a \oplus b$ can exceed a, b on <i>heavy</i> scale (pragmatically available because joint weight is stable; S&G 2017)
'Stubbornly distributive'	<i>Box A & Box B are tall (big, large, long, wide).</i> ✓Dist.: Each tall (??) Nondist.: Jointly tall because $a \oplus b$ can exceed a, b on <i>tall</i> scale (pragmatically unavailable because joint height is unstable; S&G 2017)

Table 5.2: Proposed explanation for why some adjectives are distributive, some can be understood in both ways, and some are 'stubbornly distributive'.

As with verb phrases, the distributivity potential of adjectives may appear arbitrary on the surface, so much so that one might be tempted to stipulate it. But I have argued that the behavior of these lexical items is systematically grounded in the reality that they describe.

5.5 Chapter summary

This chapter aims to explain the distributive and nondistributive understandings available to adjectives. To explain why some adjectives (*tall*) act 'stubbornly distributive' (in that their imagin-

able nondistributive understanding is not easily available) while others (*heavy*) are ‘complaisantly nondistributive’, the proposal of [Scontras & Goodman 2017](#) is endorsed: that the joint weight of a plurality of entities is more stable than its joint height, making the nondistributive understanding of *heavy* easier to coordinate on pragmatically.

As for which adjectives have an imaginable nondistributive understanding in the first place, that depends on the structure of the scale associated with the adjective, which can be captured using measurement theory. Specifically, it is argued that for a gradable adjective A to be understood nondistributively, $a \oplus b$ together must be able to exceed a and b individually on the scale invoked by A . That way, the contextual standard for what counts as A in the context can be set in such a way that $a \oplus b$ counts as A while a and b individually do not. Only adjectives with positive scales can fulfill this ordering, so only adjectives with positive scales can be understood nondistributively.

Chapter 6

Conclusion

6.1 Summary

This dissertation began from the longstanding observation that different predicates behave differently with respect to distributivity. *Smile* is distributive (true of each member of a plural subject); *meet* is nondistributive; *open the window* can go both ways. Chapter 2 argued that distributive understandings should just be contrasted with nondistributive ones, collapsing a proposed three-way semantic ambiguity between distributive, collective, and cumulative ‘readings’. Next, the bulk of the dissertation pursued two central questions:

- i (*The much-discussed compositional semantics question:*) How should inferences about distributivity be represented semantically?
- ii (*The less-discussed lexical semantics question:*) Which predicates behave in which ways, and why?

To address (i), Chapter 3 put forward a unified, fundamentally pragmatic analysis of distributivity whereby any predicate applied to a plural is true of each cell of some contextually determined cover of the subject. All inferences about distributivity are framed as inferences about which cover(s) to entertain, depending on what is known about the event or state described by the predicate.

To address (ii), Chapter 4 used a large-scale dataset to generalize the analysis of *smile*, *meet*, and *open the window* to over 1637 verb phrases. Other body / mind verb phrases act like *smile*; other multilateral verb phrases act like *meet*; causatives and incremental-object predicates act like *open the window*. Together, these patterns also indirectly explain why intransitive verbs tend to be distributive, while those built from transitives tend to allow a nondistributive understanding: because many intransitives are body / mind verbs (distributive), while many transitives are causative and / or have an incremental object (creating the potential for a nondistributive understanding).

Chapter 5 used tools from measurement theory to make predictions about adjectives, arguing that an adjective can only be understood nondistributively if it is associated with a scale that behaves ‘positively’ with respect to concatenation (*the boxes are heavy* can be nondistributive because two boxes together are heavier than each one individually). The underspecified semantics from Chapter 3 becomes explanatory when combined with a predictive analysis of which predicates are understood in which ways.

6.2 Open questions

This dissertation has made progress in seeking a predictive theory of distributivity, but it leaves many questions open. First, I have focused on determining which ways of understanding a predicate are possible (*open the window* can be distributive or nondistributive); but it is also worth investigating which ways are more preferred or frequent. Other work has shown that when a predicate can be understood in both ways, the nondistributive understanding is strongly preferred (§1.3.4 — although the reason for this preference remains open). Future work might investigate the strength of such preferences, and how that depends on the nature of the predicate — whether it involves a verb or an adjective; whether the object (if there is one) is definite or indefinite, singular or plural, count or mass, and what the object refers to. In the Distributivity Ratings Dataset, each transitive verb is tested with only one object, but it would also be interesting to test the same verb with multiple different objects: how would *open a soda* or *open a vault* compare to *open a window*, given the difficulty of opening these different objects?

Moreover, when people encounter a sentence with a plural subject, they may not settle on one way of understanding the sentence, but might entertain different options with some considered more likely than others; or might not even care whether the predicate is understood distributively or not. Future work might explore these calculations.

It will also be valuable to look beyond conjoined names towards numeral, definite, and quantified plurals (*three children, the children, some children*). Conjoined names were chosen to avoid nonmaximality (*the children smiled* may admit exceptions, whereas *Alice and Bob smiled* seems not to). But nonmaximality interacts in interesting ways with lexical semantics and pragmatics (Dowty 1987, Yoon 1996, Krifka 1996): *the reporters asked questions* may convey that only some of them did, while *the reporters were silent* suggests that all of them were (similarly: *the glasses are clean* conveys that all of them are clean, while *the glasses are dirty* may convey that only some are dirty; Yoon 1996). While there are theories modeling these ‘universal’ and ‘existential’ readings (Malamud 2012, Križ 2016, Champollion et al. to appear), it is an open question which predicates are more or less resistant to exceptions.

Most importantly, by grounding distributivity in ‘world knowledge’, this dissertation makes clear crosslinguistic predictions, which should be tested.¹ If the behavior of various predicates is tied to language-independent facts about the things they describe, then such predicates are predicted to act the same in any language: if a language has a word for *smile*, it should be distributive applied to *Alice and Bob*. But of course, different languages might lexicalize similar events in different ways, or might have different grammatical resources (conjunction morphemes, distributivity markers, syntactic effects on distributivity, and so on). Thus, even if distributivity consistently depends on ‘world knowledge’ as predicted, the view from English is unlikely to be universal.

¹For non-Indo-European work on distributivity, see for example, Choe 1987 and Joh 2008 for Korean; Ouwayda 2014 and Ouwayda 2017 for Arabic; and Lin 1998, Kratzer 2007: §7, and Xiang 2008 for Mandarin, where the literature seems to disagree on whether a predicate which could theoretically go both ways (*buy a car, eat an apple pie*) can be understood distributively in the absence of the (much-discussed, multi-functional) distributivity marker *dōu*.

6.3 Zooming out

While distributivity is quite a specialized topic, I believe that it engages with larger questions in semantics, pragmatics, and linguistics as a whole. How does the structure of reality create patterns across the lexicon used to describe it? When a sentence can describe multiple different situations, should it be analyzed as ambiguous between two different meanings, or as having one general meaning compatible with both situations (Zwicky & Sadock 1975, Link 1998a)? Should a given phenomenon should be explained in terms of grammatical knowledge, or domain-general reasoning (Grice 1989, Bar-Hillel 1971)? (Of course, an explanation invoking domain-general reasoning must be made specific — which I have tried to do here.) Most fundamentally: what counts as a satisfying explanation? Is it most important to be formally explicit, or empirically predictive?

I do not have general answers to these questions, but I would like to suggest a few lessons that one might draw from this work. It is a truism that many inferences drawn from sentences depend on ‘pragmatic reasoning’ — not just reasoning about why a speaker said one thing over another, but also reasoning about the situation described by the sentence, given what is known about the world. I would like to suggest that such ‘reasoning about the world’ has as much to tell us as ‘reasoning about the speaker’. Another lesson: when a semantic theory is refined by application to a large swath of data, I would like to suggest that it not only becomes more robust as a theory of language use, but can also serve as a resource to neighboring disciplines such as natural language processing, making semantics more useful to more people.

Finally, distributivity has traditionally been studied as a compositional semantics topic. But it is defined by the observation that different predicates act differently from one another, so I would like to suggest that it has also been a lexical semantics topic all along, and that it is illuminated when treated as one.

Appendix: Further experiment testing the Incremental Hypothesis

As corroborating evidence for the Incremental Hypothesis (§4.3.6), I conducted a followup experiment where verb phrases with incremental objects were tested in two conditions, one which indicates that the predicate should be construed as telic (*ate the pizza until it was all eaten*); and one which indicates that the predicate should be construed as atelic (*ate the pizza for awhile*). The Incremental Hypothesis predicts that the ‘telic’ condition should be much more strongly nondistributive than the ‘atelic’ condition, because it is only if the verb event is carried out on the full object that each member of the subject can carry out the verb event on a different portion of it, jointly adding up to the whole (nondistributive). If the verb event is carried out on only some arbitrary portion of the object, then each member of the subject may also carry out the verb event on an arbitrary portion of the object, so that the full predicate is true of each of them (distributive).

The stimuli for this experiment were built from a list of 18 predicates with definite objects coded as ‘incremental’ in the Distributivity Ratings Dataset (§4.3.6), representing both physical and mental actions, and various ways of incrementally affecting the object — creating it, consuming it, and covering its spatial extent. The objects chosen for these verbs were the same as those used in the Distributivity Ratings Dataset (§4.2.1), except that they were definite rather than indefinite.

1. decorate the house
2. embroider the flower

- | | |
|----------------------|------------------------------|
| 3. type the letter | 11. choreograph the dance |
| 4. copy the painting | 12. compose the song |
| 5. sew the costume | 13. read the article |
| 6. build the house | 14. write the book |
| 7. weave the basket | 15. ransack the house |
| 8. drink the beer | 16. inspect the property |
| 9. eat the pizza | 17. canvass the neighborhood |
| 10. consume the fish | 18. explore the area |

Each of these predicates was randomly assigned to one of two conditions: one telic (1), and one atelic (2). Each stimulus is followed by two subquestions just as in the Distributivity Ratings Dataset, involving both the (a) ‘each’ question and the (b) ‘together’ question with five response options.

- (1) **Telic condition:** Jessica and Thomas ate the pizza until it was all eaten.
- a. Does it follow that Jessica and Thomas *each* ate the pizza until it was all eaten?
 definitely no maybe no not sure maybe yes definitely yes
- b. Could it be that Jessica and Thomas didn’t technically *each* eat the pizza until it was all eaten, because they did so *together*?
 definitely no maybe no not sure maybe yes definitely yes
- (2) **Atelic condition:** Jessica and Thomas ate the pizza for awhile.
- a. Does it follow that Jessica and Thomas *each* ate the pizza for awhile?
 definitely no maybe no not sure maybe yes definitely yes
- b. Could it be that Jessica and Thomas didn’t technically *each* eat the pizza for awhile?

definitely no maybe no not sure maybe yes definitely yes

The goal of the experiment is to compare the two conditions (1)–(2), predicting that the ‘telic’ condition will have a lower ‘each’ rating and a higher ‘together’ rating than the ‘atelic’ condition — which would indicate that a predicate is more likely to allow a nondistributive understanding when verb event is understood to be carried out on the full object by the end of the event.

Thirty-nine self-described native English speakers participated in the experiment. They answered two ‘practice’ questions intended to convey that the ‘each’ question and the ‘together’ question both ask about whether the predicate is individually true of *each* member of the subject or not, *not* whether the members of the subject were interacting socio-spatially ‘together’ while carrying out the predicate.²

Following those practice questions, each participant saw 23 questions drawn randomly from a pool of 46 potential questions: 28 fillers and 18 ‘target’ questions built from the predicates in 1–18 — each randomly assigned to either the ‘telic’ condition as in (1), or the ‘atelic’ condition as in (2).

To test the hypothesis that the telic and atelic conditions (1)–(2) differ, I conducted a mixed-effects linear regression predicting a predicate’s ‘each’ rating as a function of its condition — ‘telic’ (1) or ‘atelic’ (2). The model allows random intercepts for each participant, attributing some of the variance to unexplained differences between individual participants. The model also used random intercepts for each predicate, taking into account differences between individual predicates; and random *slopes* for each predicate, allowing that the predicted difference between the two conditions may vary depending on the particular predicate used.

According to this model, a predicate in the atelic condition (2) is predicted to have an ‘each’ rating of 3.22, while a predicate in the telic condition (1) is predicted to have a rating of 2.63 — a sizable difference (0.61 points on a 5-point scale), and a significant one ($p < 0.001$).

²One practice question guides participants to answer ‘definitely yes’ to the question of whether two people who *smile* ‘each’ do so and ‘definitely no’ to the question of whether they might not technically ‘each’ smile because they did so ‘together’. Another practice question guides them to answer ‘definitely no’ to the question of whether two people who *carry the piano upstairs* ‘each’ do so and ‘definitely yes’ to the question of whether they might not technically ‘each’ carry the piano upstairs because they did so ‘together’.

Next, I conducted another mixed-effects linear regression with the same structure (random intercepts for every participant, random intercepts and slopes for every predicate) predicting a predicate’s ‘together’ rating as a function of its condition — ‘telic’ (1) or ‘atelic’ (2). According to this model, a predicate in the atelic condition (2) is predicted to have a ‘together’ rating of 3.14, while a predicate in the telic condition (1) is predicted to have a ‘together’ rating of 3.83 — again, a sizable effect (almost 0.70 points on a 5-point scale), and this time highly significant at $p < 0.0001$. Figure 6.1 illustrates these findings.

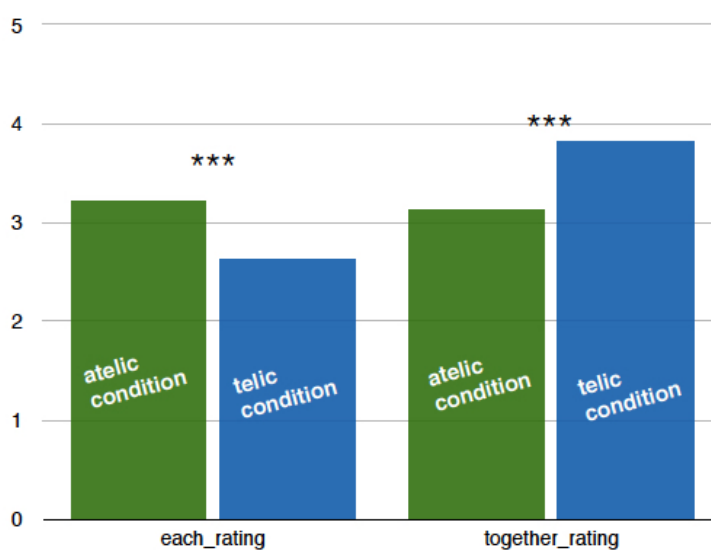


Figure 6.1: Verb phrases built from transitive verbs have systematically lower ‘each’ ratings, and systematically higher ‘together’ ratings, compared to intransitives.

These effect sizes are much larger than in the Distributivity Ratings Dataset, which I attribute to the fact that these participants were trained on how to interpret the questions while the Distributivity Ratings Dataset participants were not.

In sum, this followup experiment further demonstrates that when a predicate’s incremental object is fully affected during the event, the predicate can be understood nondistributively; whereas when its object is *not* fully affected, it may only be understood distributively (23).

The experiment also addresses some questions left open by the Distributivity Ratings Dataset. Whereas the distributivity potential of a given predicate in the Distributivity Ratings Dataset may depend on whether or not its object is actually construed as fully affected during the event — for which we have no direct data — that issue is explicitly manipulated in this experiment. As predicted, when a predicate is construed as telic, it is much more likely to allow a nondistributive understanding than when it is atelic. What matters most — even more than the particular verb or object involved — is the way the object is construed. Moreover, while the indefinite objects in the Distributivity Ratings Dataset may or may not be understood to ‘covary’ with each member of the subject, this experiment removes the potential for covariation by using definite objects. The difference between telic and atelic incremental-object predicates persists, further strengthening this finding.

Bibliography

- AKAIKE, Hirotugu (1974). A new look at the statistical model identification. *IEEE (Institute of Electrical and Electronics Engineers) Transactions on Automatic Control*, **19**(6):716–723. https://doi.org/10.1007/978-1-4612-1694-0_16.
- BACH, Emmon (1986). The algebra of events. *Linguistics and Philosophy*, **9**(1):5–16. <https://doi.org/10.1002/9780470758335.ch13>.
- BAR-HILLEL, Yehoshua (1971). Out of the pragmatic wastebasket. *Linguistic Inquiry*, **2**(3):401–406.
- BARKER, Chris (1992). Group terms in English: Representing groups as atoms. *Journal of Semantics*, **9**(1):69–93. <https://doi.org/10.1093/jos/9.1.69>.
- BARR, Dale J., Roger LEVY, Christoph SCHEEPERS, & Harry J. TILY (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3):255–278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- BARTSCH, Renate (1973). The semantics and syntax of number and numbers. In *Syntax and semantics* (edited by John Kimball), volume 2, 51–93. New York: Seminar Press.
- BARTSCH, Renate & Theo VENNEMANN (1972). The grammar of relative adjectives and comparison. *Linguistische Berichte*, **20**:19–32.
- BATES, Douglas, Reinhold KLIÉGL, Shrvan VASISHTH, & Harald BAAYEN (2015a). Parsimonious mixed models. <https://arxiv.org/abs/1506.04967>.
- BATES, Douglas, Martin MÄCHLER, Ben BOLKER, & Steve WALKER (2015b). Fitting linear

- mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1):1–48. <https://doi.org/10.18637/jss.v067.i01>.
- BECK, Sigrid & Uli SAUERLAND (2000). Cumulation is needed: A reply to Winter (2000). *Natural Language Semantics*, **8**(4):349–371. <https://doi.org/10.1023/a:1011240827230>.
- BELLETTI, Adriana & Luigi RIZZI (1988). Psych verbs and θ -theory. *Natural Language and Linguistic Theory*, **6**(3):291–352. <https://doi.org/10.1007/bf00133902>.
- BENNETT, Michael (1974). *Some extensions of a Montague fragment of English*. Ph.D. thesis, University of California, Los Angeles.
- BLOOMFIELD, Leonard (1933). *Language*. New York: Henry Holt.
- BORG, Emma (2012). Semantics without pragmatics. In *Cambridge handbook of pragmatics* (edited by Keith Allen & Kasia Jaszczolt), 513–528. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9781139022453.028>.
- BRISSON, Christine (1998). *Distributivity, maximality, and floating quantifiers*. Ph.D. thesis, Rutgers University, New Brunswick.
- BRISSON, Christine (2003). Plurals, ‘all’, and the nonuniformity of collective predication. *Linguistics and Philosophy*, **26**(2):129–184. <https://doi.org/10.1023/a:1022771705575>.
- BROOKS, Patricia J. & Martin D. BRAINE (1996). What do children know about the universal quantifiers *all* and *each*? *Cognition*, **60**(3):235–268. [https://doi.org/10.1016/0010-0277\(96\)00712-3](https://doi.org/10.1016/0010-0277(96)00712-3).
- BROWN, James Dean (2011). Likert items and scales of measurement. *SHIKEN: JALT Testing & Evaluation SIG Newsletter, Statistics Corner*, **15**(1):10–14.
- CAPPELEN, Herman & Ernest LEPORE (2005). *Intensive semantics: A defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell Publishing.
- CARIFIO, James & Rocco J. PERLA (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, **3**(3):106–116. <https://doi.org/10.3844/jssp.2007.106.116>.

- CARLSON, Greg (1998). Thematic roles and the individuation of events. In *Events and grammar* (edited by Susan Rothstein), 35–51. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-3969-4_3.
- CARSTON, Robyn (2008). Linguistic communication and the semantics/pragmatics distinction. *Synthese*, 165(3):321–345. <https://doi.org/10.1007/s11229-007-9191-8>.
- CASTAÑEDA, Hector-Neri (1967). Comments on Donald Davidson’s ‘The logical form of action sentences’. In *The logic of decision and action* (edited by Nicholas Resher), 104–112. Pittsburgh: University of Pittsburgh Press.
- CHAMPOLLION, Lucas (2010). *Parts of a whole: Distributivity as a bridge between aspect and measurement*. Ph.D. thesis, University of Pennsylvania, Philadelphia.
- CHAMPOLLION, Lucas (2016). Covert distributivity in algebraic event semantics. *Semantics and Pragmatics*, 9. <https://doi.org/10.3765/sp.9.15>.
- CHAMPOLLION, Lucas (2017). *Parts of a whole: Distributivity as a bridge between aspect and measurement*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198755128.001.0001>.
- CHAMPOLLION, Lucas (to appear). Distributivity, collectivity, and cumulativity. In *Companion to Semantics* (edited by Lisa Matthewson, Cécile Meier, Hotze Rullmann, & Thomas Ede Zimmermann). Hoboken: Wiley.
- CHAMPOLLION, Lucas, Dylan BUMFORD, & Robert HENDERSON (to appear). Donkeys under discussion. *Semantics and Pragmatics*.
- CHAMPOLLION, Lucas & Manfred KRIFKA (2015). Mereology. In *Cambridge handbook of semantics* (edited by Maria Aloni & Paul Dekker), 369–388. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9781139236157.014>.
- CHIERCHIA, Gennaro (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In *Structures and beyond: The cartography of syntactic structures* (edited by Adriana Belletti), volume 3, 39–103. New York: Oxford University Press.

- CHOE, Jae-Woong (1987). *Anti-quantifiers and a theory of distributivity*. Ph.D. thesis, University of Massachusetts, Amherst.
- CHOMSKY, Noam (1973). Conditions on transformations. In *A Festschrift for Morris Halle* (edited by Stephen Anderson & Paul Kiparsky), 232–286. New York: Holt, Rinehart & Winston.
- CHOMSKY, Noam & MORRIS HALLE (1968). *The sound pattern of English*. New York: Harper & Row.
- CONDORAVDI, Cleo & JEAN MARK GAWRON (1996). The context-dependency of implicit arguments. In *Quantifiers, deduction, and context* (edited by Makoto Kanazawa, Christopher Piñón, & Henriëtte de Swart). Stanford: CSLI (Center for the Study of Language and Information) Publications.
- COPLEY, Bridget & PHILLIP WOLFF (2014). Theories of causation should inform linguistic theory and vice versa. In *Causation in grammatical structures* (edited by Bridget Copley & Fabienne Martin), 11–57. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199672073.003.0002>.
- CRESSWELL, Maxwell John (1976). The semantics of degree. In *Montague grammar* (edited by Barbara Partee), 261–292. New York: Academic Press. <https://doi.org/10.1016/B978-0-12-545850-4.50015-7>.
- CROFT, William (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- CROFT, William (2012). *Verbs: Aspect and causal structure*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199248582.001.0001>.
- CRUSE, D. Alan (1976). Three classes of antonym in English. *Lingua*, 38(3-4):281–292. [https://doi.org/10.1016/0024-3841\(76\)90015-2](https://doi.org/10.1016/0024-3841(76)90015-2).
- DALRYMPLE, Mary, MAKOTO KANAZAWA, YOOKYUNG KIM, SAM MCHOMBO, & STANLEY PETERS (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21(2):159–210. <https://doi.org/10.1023/a:1005330227480>.
- DALRYMPLE, Mary, MAKOTO KANAZAWA, SAM MCHOMBO, & STANLEY PETERS (1994). What do

- reciprocals mean? In *Semantics and Linguistic Theory (SALT)* (edited by Mandy Harvey & Lynn Santelmann), volume 4, 61–78. Ithaca: Cornell Linguistics Circle. <https://doi.org/10.3765/salt.v4i0.2466>.
- DAVIDSON, Donald (1967). The logical form of action sentences. In *The logic of decision and action* (edited by Nicholas Rescher), 81–95. Pittsburgh: University of Pittsburgh Press. <https://doi.org/10.1093/0199246270.003.0006>.
- DAVIES, Mark (2008). The Corpus of Contemporary American English. 450 million words, 1990–present. Available online at <http://corpus.byu.edu/coca/>. Provo: Brigham Young University. <https://doi.org/10.1075/ijcl.14.2.02dav>.
- DELANCEY, Scott (1991). Event construal and case role assignment. In *Berkeley Linguistics Society (BLS)* (edited by Laurel A. Sutton, Christopher Johnson, & Ruth Shields), volume 17, 338–353. Ann Arbor: Braun-Brumfield, Inc. <https://doi.org/10.3765/bls.v17i0.1610>.
- DIXON, R. M. W. (1979). Ergativity. *Language*, 55(1):59–138. <https://doi.org/10.2307/412519>.
- DIXON, R. M. W. (1982). *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110822939>.
- DIXON, R. M. W. (2004). Adjective classes in typological perspective. In *Adjective classes: A cross-linguistic typology* (edited by R. M. W. Dixon & Alexandra Y. Aikhenvald), 1–49. Oxford: Oxford University Press.
- DOBROVIE-SORIN, Carmen, Emilia ELLSIEPEN, & Barbara HEMFORTH (2016). Why are distributive readings dispreferred? In *Romance languages and linguistic theory 10: Selected papers from 'Going Romance' 28* (edited by Ernestina Carrilho, Alexandra Fiéis, Maria Lobo, & Sandra Pereira), 83–102. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.10.05dob>.
- DOTLAČIL, Jakub (2010). *Anaphora and distributivity: A study of same, different, reciprocals and others*. Ph.D. thesis, Utrecht University, Utrecht.

- DOWTY, David R. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel. <https://doi.org/10.1007/978-94-009-9473-7>.
- DOWTY, David R. (1987). A note on collective predicates, distributive predicates, and ‘all’. In *Eastern States Conference on Linguistics (ESCOL)* (edited by Fred Marshall, Ann Miller, & Zheng-sheng Zhang), volume 3, 97–116. Columbus: The Ohio State University.
- DOWTY, David R. (1991). Thematic proto-roles and argument selection. *Language*, **67**(3):547–619. <https://doi.org/10.1353/lan.1991.0021>.
- FILLMORE, Charles J. (1969). Types of lexical information. In *Studies in syntax and semantics* (edited by Ferenc Kiefer), 109–137. Dordrecht: Reidel. https://doi.org/10.1007/978-94-010-1707-7_6.
- FILLMORE, Charles J. (1970). The grammar of hitting and breaking. In *Readings in English transformational grammar* (edited by Roderick Jacobs & Peter Rosenbaum), 120–33. Washington: Georgetown University Press.
- FRAZIER, Lyn, Jeremy M. PACTH, & Keith RAYNER (1999). Taking on semantic commitments, II: Collective versus distributive readings. *Cognition*, **70**(1):87–104. [https://doi.org/10.1016/s0010-0277\(99\)00002-5](https://doi.org/10.1016/s0010-0277(99)00002-5).
- GAMUT, L.T.F. (1991). *Logic, language, and meaning*, volume 1: Introduction to logic. Chicago: University of Chicago Press. ‘L.T.F. Gamut’ is a collective pseudonym for Johan van Benthem, Jeroen Groenendijk, Dick de Jongh, Martin Stockhof, and Henk Verkuyl.
- GELMAN, Andrew & Jennifer HILL (2007). *Data analysis using regression and multi-level/hierarchical models*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>.
- GILLON, Brendan S. (1987). The readings of plural noun phrases in English. *Linguistics and Philosophy*, **10**(2):199–219. <https://doi.org/10.1007/bf00584318>.
- GILLON, Brendan S. (1990). Plural noun phrases and their readings: A reply to Lasersohn. *Linguistics and Philosophy*, **13**(4):477–485. <https://doi.org/10.1007/bf00630751>.
- GLASS, Lelia (2017). Exploring the relation between argument structure and distributivity. In

- Berkeley Linguistics Society (BLS)* (edited by Julia Nee, Margaret Cychosz, Dmetri Hayes, Tyler Lau, & Emily Ramirez), volume 43, 95–119. Berkeley: eScholarship.
- GLASS, Lelia (2018). Deriving the distributivity potential of adjectives via measurement theory. In *Linguistic Society of America (LSA)* (edited by Patrick Ferrell), volume 3. Washington, D.C.: Linguistic Society of America. <https://doi.org/10.3765/plsa.v3i1.4343>.
- GLASS, Lelia & Nanjiang JIANG (2017). Distributivity ratings dataset. Available online through the Open Science Framework at <https://doi.org/10.17605/osf.io/8953e>. Stanford University.
- GRICE, H. P. (1989). Logic and conversation. In *Studies in the way of words*. Cambridge: Harvard University Press. Originally published as Grice (1975), in *Syntax and semantics* (edited by Peter Cole and Jerry L. Morgan), volume 3: Speech acts, 41–58. New York: Academic Press.
- GROPEN, Jess, Steven PINKER, Michelle HOLLANDER, & Richard GOLDBERG (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, **41**(1):153–195. [https://doi.org/10.1016/0010-0277\(91\)90035-3](https://doi.org/10.1016/0010-0277(91)90035-3).
- GUÉRON, Jacqueline (2006). Inalienable possession. In *Blackwell companion to syntax* (edited by Martin Everaert & Henk van Riemsdijk), volume 1, 589–638. Oxford: Blackwell Publishing.
- HAGOORT, Peter, Lea HALD, Marcel BASTIAANSEN, & Karl Magnus PETERSSON (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, **304**(5669):438–441. <https://doi.org/10.1126/science.1095455>.
- HARLEY, Heidi & Rolf NOYER (2000). Formal versus encyclopedic properties of vocabulary: Evidence from nominalisations. In *The lexicon-encyclopedia interface* (edited by Bert Peters), 246–323. New York: Elsevier.
- HIGGINBOTHAM, James (1981). Reciprocal interpretation. *Journal of Linguistic Research*, **1**(3):97–117.
- HIGGINBOTHAM, James (1985). On semantics. *Linguistic Inquiry*, **16**(4):547–594.
- HOBBS, Jerry R (1987). World knowledge and word meaning. In *Workshop on theoretical issues*

- in natural language processing*, 20–27. Stroudsburg: Association for Computational Linguistics.
<https://doi.org/10.1515/9783110226614.740>.
- HOEKSEMA, Jack (1988). The semantics of non-boolean ‘and’. *Journal of Semantics*, 6(1):19–40.
<https://doi.org/10.1093/jos/6.1.19>.
- HOPPER, Paul J & Sandra A THOMPSON (1980). Transitivity in grammar and discourse. *Language*, 56(2):251–299. <https://doi.org/10.2307/413757>.
- HORNBY, Albert Sydney, Sally WEHMEIER, & Michael ASHBY (1995). *Oxford advanced learner’s dictionary*. Oxford: Oxford University Press, 9 edition.
- HUME, David (1748). *Enquiries concerning the human understanding: And concerning the principles of morals*. Oxford: Clarendon Press (1902). <https://doi.org/10.1093/oseo/instance.00046349>.
- HYDE, Dominic & Diana RAFFMAN (2014). Sorites paradox. In *The Stanford Encyclopedia of Philosophy* (edited by Edward N. Zalta). Stanford: Stanford University Metaphysics Research Lab. <https://plato.stanford.edu/archives/sum2018/entries/sorites-paradox/>.
- JACKENDOFF, Ray (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory*, 14(2):305–354. <https://doi.org/10.1007/bf00133686>.
- JOH, Yoon-kyoung (2008). *Plurality and distributivity*. Ph.D. thesis, University of Pennsylvania, Philadelphia.
- KAPLAN, David (1977). On the logic of demonstratives. *Journal of Philosophical Logic*, 8(1). <https://doi.org/10.1007/bf00258420>.
- KATZ, Jerrold J. & Jerry A. FODOR (1963). The structure of a semantic theory. *Language*, 39(2):170–210. <https://doi.org/10.2307/411200>.
- KAUP, Barbara, Stephanie KELTER, & Christopher HABEL (2002). Representing referents of plural expressions and resolving plural anaphors. *Language and Cognitive Processes*, 17(4):399–467. <https://doi.org/10.1080/01690960143000272>.

- KENNEDY, Christopher (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland. <https://doi.org/10.4324/9780203055458>.
- KENNEDY, Christopher (2001). Polar opposition and the ontology of ‘degrees’. *Linguistics and Philosophy*, **24**(1):33–70. <https://doi.org/10.1023/a:1005668525906>.
- KENNEDY, Christopher (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, **30**(1):1–45. <https://doi.org/10.1007/s10988-006-9008-0>.
- KENNEDY, Christopher & Louise MCNALLY (2005). Scale structure, degree modification, and the semantic typology of gradable predicates. *Language*, **81**(2):345–381. <https://doi.org/10.1353/lan.2005.0071>.
- KLEIN, Ewan (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, **1**(4):1–45. <https://doi.org/10.1007/BF00351812>.
- KOPTJEVSKAJA-TAMM, Maria (2011). ‘It’s boiling hot’: On the structure of the linguistic temperature domain across languages. In *Rahmen des Sprechens: Beiträge zur Valenztheorie, Varietätenlinguistik, Kognitiven und Historischen Semantik (Frame of speech: Contributions to valence theory, variationist linguistics, cognitive and historical semantics)* (edited by Sarah Dessì Schmid, Ulrich Detges, Paul Gévaudan, Wiltrud Mihatsch, & Richard Waltereit), 393–410. Tübingen: Narr.
- KOPTJEVSKAJA-TAMM, Maria & Ekaterina V. RAKHILINA (2006). ‘Some like it hot’: On semantics of temperature adjectives in Russian and Swedish. *STUF (Sprachtypologie und Universalienforschung / Language typology and universals), a special issue on the lexicon in a typological and contrastive perspective*, **59**(3):253–269. <https://doi.org/10.1524/stuf.2006.59.3.253>.
- KRANTZ, David H, R. Duncan LUCE, Patrick SUPPES, & Amos TVERSKY (1971). *Foundations of measurement*. San Diego: Academic Press.
- KRATZER, Angelika (2007). On the plurality of verbs. In *Event structures in linguistic form and*

- interpretation* (edited by Johannes Dölling, Tatjana Heyde-Zybatow, & Martin Schäfer), 269–300. Berlin: De Gruyter. <https://doi.org/10.1515/9783110925449.269>.
- KRIFKA, Manfred (1989). Nominal reference, temporal constitution and quantification in event semantics. In *Semantics and contextual expression* (edited by Renate Bartsch, Johan van Benthem, & Peter van Emde Boas), 75–115. Berlin: De Gruyter.
- KRIFKA, Manfred (1992). Thematic relations as links between nominal reference and temporal constitution. In *Lexical Matters* (edited by Ivan A. Sag & Anna Szabolcsi), 29–54. Stanford: CSLI (Center for the Study of Language and Information) Publications.
- KRIFKA, Manfred (1996). Pragmatic strengthening in plural predications and donkey sentences. In *Semantics and Linguistic Theory (SALT)* (edited by Teresa Galloway & Justin Spence), volume 6, 136–153. Ithaca: Cornell Linguistics Circle. <https://doi.org/10.3765/salt.v6i0.2769>.
- KRIŽ, Manuel (2016). Homogeneity, non-maximality, and ‘all’. *Journal of Semantics*, **3**(33):493–539. <https://doi.org/10.1093/jos/ffv006>.
- KROCH, Anthony S. (1974). *The semantics of scope in English*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.
- KRUITWAGEN, Imke, Eva B. POORTMAN, & Yoad WINTER (2017). Reciprocal verbs as collective predicate concepts. In *North East Linguistic Society (NELS)* (edited by Andrew Lamont & Kateřina Tetzloff), volume 47. Amherst: Graduate Linguistics Student Association of the University of Massachusetts.
- LANDMAN, Fred (1989a). Groups, I. *Linguistics and Philosophy*, **12**(5):559–605. <https://doi.org/10.1007/bf00627774>.
- LANDMAN, Fred (1989b). Groups, II. *Linguistics and Philosophy*, **12**(6):723–744. <https://doi.org/10.1007/bf00632603>.
- LANDMAN, Fred (1996). Plurality. In *The handbook of contemporary semantic theory* (edited by Shalom Lappin), 425–457. Oxford: Blackwell Publishing.

- LANDMAN, Fred (2000). *Events and plurality: The Jerusalem lectures*. Dordrecht: Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-011-4359-2>.
- LANGACKER, Ronald W (1987). *Foundations of cognitive grammar: Theoretical prerequisites*, volume 1. Stanford: Stanford University Press.
- LASERSOHN, Peter (1988). *A semantics for groups and events*. Ph.D. thesis, The Ohio State University, Columbus.
- LASERSOHN, Peter (1989). On the readings of plural noun phrases. *Linguistic Inquiry*, **20**(1):130–134.
- LASERSOHN, Peter (1990a). Group action and spatio-temporal proximity. *Linguistics and Philosophy*, **13**(2):179–206. <https://doi.org/10.1007/bf00630733>.
- LASERSOHN, Peter (1990b). *A semantics for groups and events*. New York: Garland.
- LASERSOHN, Peter (1993). Lexical distributivity and implicit arguments. In *Semantics and Linguistic Theory (SALT)* (edited by Utpal Lahiri & Adam Wyner), volume 3, 145–161. Ithaca: Cornell University. <https://doi.org/10.3765/salt.v3i0.2751>.
- LASERSOHN, Peter (1995). *Plurality, conjunction and events*. Dordrecht: Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-015-8581-1>.
- LASERSOHN, Peter (1998a). Events in the semantics of collectivizing adverbials. In *Events and grammar* (edited by Susan Rothstein), 273–292. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-3969-4_11.
- LASERSOHN, Peter (1998b). Generalized distributivity operators. *Linguistics and Philosophy*, **21**(1):83–93. <https://doi.org/10.1023/a:1005317815339>.
- LASERSOHN, Peter (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, **28**(6):643–686. <https://doi.org/10.1007/s10988-005-0596-x>.
- LASERSOHN, Peter (2011). Mass nouns and plurals. In *Semantics: An international handbook of natural language meaning* (edited by Klaus von Heusinger, Claudia Maienborn, &

- Paul Portner), volume 2, 1131–1153. Berlin: De Gruyter. <https://doi.org/10.1515/9783110255072.1131>.
- LASSITER, Daniel (2011). *Measurement and modality*. Ph.D. thesis, New York University, New York.
- LASSITER, Daniel (2017). *Graded modality: Qualitative and quantitative perspectives*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198701347.001.0001>.
- LASSITER, Daniel & Noah D. GOODMAN (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. In *Semantics and Linguistic Theory (SALT)* (edited by Todd Snider), volume 23, 587–610. Washington, D.C.: Linguistic Society of America. <https://doi.org/10.3765/salt.v23i0.2658>.
- LEHRER, Adrienne & Keith LEHRER (1982). Antonymy. *Linguistics and Philosophy*, 5(4):483–501. <https://doi.org/10.1007/bf00355584>.
- LEVIN, Beth (1993). *English verb classes and alternations*. Chicago: University of Chicago Press.
- LEVIN, Beth (2000). Aspect, lexical semantic properties, and argument expression. In *Berkeley Linguistics Society (BLS)* (edited by Lisa J. Conathan, Jeff Good, Darya Kavitskaya, Alyssa B. Wulf, & Alan C. L. Yu), volume 26, 413–429. Ann Arbor: Sheridan Books. <https://doi.org/10.3765/bls.v26i1.1129>.
- LEVIN, Beth & Malka RAPAPORT HOVAV (2014). Manner and result: A view from ‘clean’. In *Language description informed by theory* (edited by Rob Pensalfini, Myfany Turpin, & Diana Guillemin), 337–357. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.147.14lev>.
- LEVIN, Beth & Malka RAPPAPORT HOVAV (2005). *Argument realization*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511610479>.
- LEVINSON, Stephen C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- LEVINSON, Stephen C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: MIT Press.

- LEWIS, David (1969). *Convention*. Cambridge: Harvard University Press. <https://doi.org/10.1002/9780470693711>.
- LEWIS, David (1973). Causation. *The Journal of Philosophy*, **70**(17):556–567. <https://doi.org/10.2307/2025310>.
- LIEBER, Rochelle (1980). *On the organization of the lexicon*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.
- LIN, Jo-Wang (1998). Distributivity in Chinese and its implications. *Natural Language Semantics*, **6**(2):201–243. <https://doi.org/10.1023/a:1008299031574>.
- LINK, Godehard (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, Use and Interpretation of Language* (edited by Rainer Bäuerle, Christoph Schwarze, & Arnim von Stechow), 127–146. Berlin: De Gruyter. <https://doi.org/10.1515/9783110852820.302>.
- LINK, Godehard (1991). Plural. In *Semantik: Ein internationales Handbuch Handbuch der zeitgenoessischen Forschung [Semantics: An international handbook of contemporary research]* (edited by Arnim von Stechow & Dieter Wunderlich). Berlin: De Gruyter. <https://doi.org/10.1515/9783110126969.6.418>.
- LINK, Godehard (1998a). *Algebraic semantics in language and philosophy*. Stanford: CSLI (Center for the Study of Language and Information) Publications.
- LINK, Godehard (1998b). Ten years of research on plurals – Where do we stand? In *Plurality and quantification* (edited by Fritz Hamm & Erhard Hinrichs), Studies in Linguistics and Philosophy, 19–54. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-2706-8_2.
- LYUTIKOVA, Ekaterina & Sergei TATEVOSOV (2014). Causativization and event structure. In *Causation in grammatical structures* (edited by Bridget Copley & Fabienne Martin), Oxford Studies in Theoretical Linguistics 52, 279–327. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199672073.003.0011>.
- MADOR-HAIM, Sela & Yoad WINTER (2015). Far from obvious: The semantics of locative

- indefinites. *Linguistics and Philosophy*, **38**(5):437–476. <https://doi.org/10.1007/s10988-015-9175-y>.
- MAGRI, Giorgio (2012). Collective nouns without groups. In *Israeli Association for Theoretical Linguistics (IATL)* (edited by Evan Cohen), volume 27, 183–202. Cambridge: MIT Working Papers in Linguistics.
- MALAMUD, Sophia (2012). The meaning of plural definites: A decision-theoretic approach. *Semantics and Pragmatics*, **5**. <https://doi.org/10.3765/sp.5.3>.
- MALAMUD, Sophia A. (2006). (Non)maximality and distributivity: A decision theory approach. In *Semantics and Linguistic Theory (SALT)* (edited by Masayuki Gibson & Jonathan Howell), volume 16, 120–137. Ithaca: Cornell Linguistics Circle. <https://doi.org/10.3765/salt.v16i0.2943>.
- MALDONADO, Mora, Emmanuel CHEMLA, & Benjamin SPECTOR (2017). Priming plural ambiguities. *Journal of Memory and Language*, **95**:89 – 101. <https://doi.org/10.1016/j.jml.2017.02.002>.
- MCCAWLEY, James D. (1968). The role of semantics in a grammar. In *Universals in Linguistic Theory* (edited by Emmon Bach & Richard Harms), 124–169. New York: Holt, Rinehart and Winston.
- MCNALLY, Louise (1997). *A semantics for the English existential construction*. New York: Garland.
- MCNALLY, Louise (2013). Semantics and pragmatics. *Wiley Interdisciplinary Reviews: Cognitive Science*, **4**(3):285–297. <https://doi.org/10.1002/wcs.1227>.
- MOLTMANN, Friederike (1997). *Parts and wholes in semantics*. Oxford: Oxford University Press.
- MOLTMANN, Friederike (2004). The semantics of ‘together’. *Natural Language Semantics*, **12**(4):289–318. <https://doi.org/10.1007/s11050-004-6453-6>.
- MORRIS, Charles William (1938). Foundations of the theory of signs. In *International encyclopedia of unified science* (edited by Otto Neurath, Rudolf Carnap, & Charles Morris), 1–59. Chicago: University of Chicago Press.

- MUEHLEISEN, Victoria Lynn (1997). *Antonymy and semantic range in English*. Ph.D. thesis, Northwestern University, Evanston.
- NAIGLES, Letitia (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17(2):357–374. <https://doi.org/10.1017/s0305000900013817>.
- NAIGLES, Letitia & Edward T KAKO (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, 64(6):1665–1687. <https://doi.org/10.2307/1131462>.
- NEELEMAN, Ad & Hans VAN DE KOOT (2012). The linguistic expression of causation. In *The theta system: Argument structure at the interface* (edited by Martin Everaert & Marijana Marelj), 20–51. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199602513.003.0002>.
- NOUWEN, Rick (2015). Plurality. In *Cambridge handbook of semantics* (edited by Paul Dekker & Maria Aloni), 267–284. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9781139236157.010>.
- OUWAYDA, Sarah (2014). *Where number lies: Plural marking, numerals, and the collective-distributive distinction*. Ph.D. thesis, University of Southern California, Los Angeles.
- OUWAYDA, Sarah (2017). On the DP dependence of collective interpretation with numerals. *Natural Language Semantics*, 25(4):263–314. <https://doi.org/10.1007/s11050-017-9136-9>.
- PAGLIARINI, Elena, Gaetano FIORIN, & Jakub DOTLAČIL (2012). The acquisition of distributivity in pluralities. In *Boston University Conference on Language Development (BUCLD)*, volume 36, 387–399. Somerville: Cascadilla Press.
- PARSONS, Terence (1990). *Events in the semantics of English: A study in subatomic semantics*. Cambridge: MIT Press.
- PEETERS, Bert (2000). Setting the scene: Some recent milestones in the lexicon-encyclopedia debate. In *The lexicon-encyclopedia interface* (edited by Bert Peeters), 1–52. New York: Elsevier.
- POORTMAN, Eva B., Marijn E. STRUIKSMA, Nir KEREM, Naama FRIEDMANN, & Yoad WINTER

- (2018). Reciprocal expressions and the Maximal Typicality Hypothesis. *Glossa*, 3(1). <https://doi.org/10.5334/gjgl.180>.
- POTTS, Christopher (2008). Interpretive economy, Schelling points, and evolutionary stability. Manuscript, University of Massachusetts, Amherst.
- POTTS, Christopher, Daniel LASSITER, Roger LEVY, & Michael C. FRANK (2016). Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics*, 33(4):755–802. <https://doi.org/10.1093/jos/ffv012>.
- PUSTEJOVSKY, James (1995). *The generative lexicon*. Cambridge: MIT Press.
- QUINE, Willard Van Orman (1960). *Word and object*. Cambridge: MIT press.
- RAPPAPORT HOVAV, Malka (2008). Lexicalized meaning and the internal structure of events. In *Theoretical and crosslinguistic approaches to the semantics of aspect* (edited by Susan Rothstein), 13–42. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.110.03hov>.
- RAPPAPORT HOVAV, Malka & Beth LEVIN (2010). Reflections on manner/result complementarity. In *Syntax, lexical semantics, and event structure* (edited by Edit Doron, Malka Rappaport Hovav, & Ivy Sichel). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199544325.001.0001>.
- REINHART, Tanya (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy*, 20(4):335–397. <https://doi.org/10.1023/a:1005349801431>.
- ROBERTS, Craige (1987). *Modal subordination, anaphora, and distributivity*. Ph.D. thesis, University of Massachusetts, Amherst.
- ROTHSTEIN, Susan (2001). What are incremental themes? In *Papers on predicative constructions* (edited by Gerhard Jäger, Anatoliy Strigin, Chris Wilder, & Ning Zhang), volume 22 of *ZAS Papers in Linguistics*, 139–157. Berlin: Zentrum für Allgemeine Sprachwissenschaft [Center for General Linguistics].

- ROTHSTEIN, Susan (2004). *Structuring events: A study in the semantics of aspect*. Oxford: Blackwell Publishing. <https://doi.org/10.1002/9780470759127>.
- ROTHSTEIN, Susan (2012). Another look at accomplishments and incrementality. In *Telicity, change, and state: A cross-categorial view of event structure* (edited by Violeta Demonte & Louise McNally), 60–102. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199693498.003.0003>.
- ROTHSTEIN, Carmen & Yoad WINTER (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics*, **12**(3):259–288. <https://doi.org/10.1023/b:nals.0000034517.56898.9a>.
- RULLMANN, Hotze (1995). *Maximality in the semantics of wh-constructions*. Ph.D. thesis, University of Massachusetts, Amherst.
- RUSSELL, Benjamin (2006). Against grammatical computation of scalar implicatures. *Journal of Semantics*, **23**(4):361–382. <https://doi.org/10.1093/jos/ffl008>.
- SAPIR, Edward (1944). Grading: A study in semantics. *Philosophy of Science*, **11**(2):93–116. <https://doi.org/10.1086/286828>.
- SASSOON, Galit Weidman (2007). *Vagueness, gradability and typicality: A comprehensive semantic analysis*. Ph.D. thesis, Tel Aviv University, Tel Aviv.
- SASSOON, Galit Weidman (2010). Measurement theory in linguistics. *Synthese*, **174**(1):151–180. <https://doi.org/10.1007/s11229-009-9687-5>.
- DE SAUSSURE, Ferdinand (1916). *Cours de linguistique générale*. Paris: Payot.
- SCHA, Remko (1981). Distributive, collective and cumulative quantification. In *Formal methods in the study of language* (edited by Jeroen Groenendijk & Martin Stockhof), 483–512. Amsterdam: Mathematical Center Tracts. <https://doi.org/10.1515/9783110867602.131>.
- SCHEIN, Barry (1986). *Event logic and the interpretation of plurals*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.
- SCHEIN, Barry (1993). *Plurals and events*. Cambridge: MIT Press.
- SCHWARZSCHILD, Roger (1992). ‘Together’ as a non-distributivity marker. In *Proceedings of*

- the Amsterdam Colloquium* (edited by Paul Dekker & Martin Stockhof), volume 8. Amsterdam: ILLC.
- SCHWARZSCHILD, Roger (1994). Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics*, 2(3):201–248. <https://doi.org/10.1007/bf01256743>.
- SCHWARZSCHILD, Roger (1996). *Pluralities*. Dordrecht: Kluwer Academic Publishers.
- SCHWARZSCHILD, Roger (2002). The grammar of measurement. In *Semantics and Linguistic Theory (SALT)* (edited by Brendan Jackson), volume 12, 225–245. Ithaca: Cornell Linguistics Circle. <https://doi.org/10.3765/salt.v12i0.2870>.
- SCHWARZSCHILD, Roger (2006). The role of dimensions in the syntax of noun phrases. *Syntax*, 9(1):67–110. <https://doi.org/10.1111/j.1467-9612.2006.00083.x>.
- SCHWARZSCHILD, Roger (2011). Stubborn distributivity, multiparticipant nouns and the count/mass distinction. In *North East Linguistic Society (NELS)* (edited by Suzi Lima, Kevin Mullin, & Brian Smith), volume 39. Amherst: Graduate Linguistics Student Association of the University of Massachusetts.
- SCONTRAS, Gregory & Noah GOODMAN (2017). Resolving uncertainty in plural predication. *Cognition*, 168:294–311. <https://doi.org/10.1016/j.cognition.2017.07.002>.
- SEARLE, John R. (1978). Literal meaning. *Erkenntnis*, 13(1):207–224. <https://doi.org/10.1007/bf00160894>.
- SEUREN, Pieter A. M. (1973). The comparative. In *Generative grammar in Europe* (edited by Ferenc Keifer & Nicolas Ruwet), 528–563. Dordrecht: Reidel. https://doi.org/10.1007/978-94-010-2503-4_22.
- SEUREN, Pieter A. M. (1978). The structure and selection of positive and negative gradable adjectives. In *Parasession on the lexicon, Chicago Linguistic Society (CLS)* (edited by Donka Farkas, Wesley M. Jacobsen, & Karol W. Todrys), volume 14, 336–346. Chicago: Chicago Linguistic Society.
- SHIBATANI, Masayoshi (1973). Semantics of Japanese causativization. *Foundations of language*, 9(3):327–373.

- SILONI, Tal (2012). Reciprocal verbs and symmetry. *Natural Language and Linguistic Theory*, **30**(1):261–320. <https://doi.org/10.1007/s11049-011-9144-2>.
- SMITH, Carlota S. (1970). Jespersen's 'move and change' class and causative verbs in English. In *Linguistic and literary studies in honor of Archibald A. Hill* (edited by Mohammad A. Jazayery, Edgar C. Polomé, & Werner Winter), volume 2: Descriptive Linguistics, 101–109. The Hague: Mouton. <https://doi.org/10.1515/9783110800432.101>.
- SMOLLETT, Rebecca (2005). Quantized direct objects don't delimit after all. In *Perspectives on aspect* (edited by Henk J. Verkuyl, Henriëtte de Swart, & Angeliek van Hout), 41–59. Dordrecht: Springer. https://doi.org/10.1007/1-4020-3232-3_3.
- SOLT, Stephanie (2015). Measurement scales in natural language. *Language and Linguistics Compass*, **9**(1):14–32. <https://doi.org/10.1111/lnc3.12101>.
- VON STECHOW, Arnim (1984). Comparing semantic theories of comparison. *Journal of Semantics*, **3**(1-2):1–77. <https://doi.org/10.1093/jos/3.1-2.1>.
- STEVENS, Stanley Smith (1946). On the theory of scales of measurement. *Science*, **103**(2684):688–680. <https://doi.org/10.1126/science.103.2684.677>.
- SUPPES, Patrick & Joseph L. ZINNES (1962). *Basic measurement theory*. Stanford: Institute for Mathematical Studies in the Social Sciences.
- SYRETT, Kristen (2015). Mapping properties to individuals in language acquisition. In *Boston University Conference on Language Development (BUCLD)*, volume 39, 398–410. Somerville: Cascadilla Press.
- SYRETT, Kristen & Julien MUSOLINO (2013). Collectivity, distributivity, and the interpretation of plural numerical expressions in child and adult language. *Language Acquisition*, **20**(4):259–291. <https://doi.org/10.1080/10489223.2013.828060>.
- SYRETT, Kristen & Julien MUSOLINO (2016). All together now: Disentangling semantics and pragmatics with 'together' in child and adult language. *Language Acquisition*, **23**(2):175–197. <https://doi.org/10.1080/10489223.2015.1067319>.
- SZABÓ, Zoltán Gendler (2008). The distinction between semantics and pragmatics. In *The*

- Oxford handbook of philosophy of language* (edited by Ernest Lepore & Barry C. Smith), 361 – 389. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199552238.003.0017>.
- TAYLOR, Kenneth A (2001). Sex, breakfast, and descriptus interruptus. *Synthese*, **128**(1-2):45–61.
- TENNY, Carol L. (1987). *Grammaticalizing aspect and affectedness*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.
- TENNY, Carol L. (1994). *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-011-1150-8>.
- UNGER, Peter (1978). *Ignorance: A case for scepticism*. Oxford: Clarendon Press. <https://doi.org/10.1093/0198244177.001.0001>.
- VENDLER, Zeno (1967). *Linguistics in philosophy*. Ithaca: Cornell University Press.
- VERKUYL, Henk (1972). *On the compositional nature of the aspects*. Dordrecht: Reidel. <https://doi.org/10.1007/978-94-017-2478-4>.
- VERKUYL, Henk (1994). Distributivity and collectivity: A couple at odds. In *Dynamics, polarity and quantification* (edited by Makoto Kanazawa & Christopher J. Piñón), 49–80. Stanford: CSLI (Center for the Study of Language and Information) Publications.
- VERKUYL, Henk J. & Jaap VAN DER DOES (1996). The semantics of plural noun phrases. In *Quantifiers, logic, and language* (edited by Jaap van der Does & Jan van Eijck). Stanford: CSLI (Center for the Study of Language and Information) Publications.
- DE VRIES, Hanna (2015). *Shifting sets, hidden atoms: The semantics of distributivity, plurality and animacy*. Ph.D. thesis, Utrecht University, Utrecht.
- DE VRIES, Hanna (2017). Two kinds of distributivity. *Natural Language Semantics*, **25**(2):173–197. <https://doi.org/10.1007/s11050-017-9133-z>.
- WINTER, Bodo (2013). Linear models and linear mixed effects models in R with linguistic applications. University of California, Merced, tutorial, available at <http://arxiv.org/pdf/1308.5499.pdf>.
- WINTER, Yoad (1996). What does the Strongest Meaning Hypothesis mean? In *Semantics and*

- Linguistic Theory (SALT)* (edited by Teresa Galloway & Justin Spence), volume 6, 295–310. Ithaca: Cornell Linguistics Circle. <https://doi.org/10.3765/salt.v6i0.2776>.
- WINTER, Yoad (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy*, 20(4):399–467. <https://doi.org/10.1023/a:1005354323136>.
- WINTER, Yoad (2000). Distributivity and dependency. *Natural Language Semantics*, 8(1):27–69. <https://doi.org/10.1023/a:1008313715103>.
- WINTER, Yoad (2001a). *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge: MIT Press.
- WINTER, Yoad (2001b). Plural predication and the Strongest Meaning Hypothesis. *Journal of Semantics*, 18(4):333–365. <https://doi.org/10.1093/jos/18.4.333>.
- WINTER, Yoad (2002). Atoms and sets: A characterization of semantic number. *Linguistic Inquiry*, 33(3):493–505. <https://doi.org/10.1162/002438902760168581>.
- WINTER, Yoad (2018). Symmetric predicates and the semantics of reciprocal alternations. *Semantics and Pragmatics*, 11. <https://doi.org/10.3765/sp.11.1>.
- WINTER, Yoad & Remko SCHA (2015). Plurals. In *Handbook of contemporary semantic theory* (edited by Shalom Lappin & Chris Fox), 77–113. New York: Wiley, 2 edition. <https://doi.org/10.1002/9781118882139.ch3>.
- XIANG, Ming (2008). Plurality, maximality and scalar inferences: A case study of Mandarin ‘dou’. *Journal of East Asian Linguistics*, 17(3):227–245. <https://doi.org/10.1007/s10831-008-9025-9>.
- YOON, Youngeun (1996). Total and partial predicates and the weak and strong interpretations. *Natural Language Semantics*, 4(3):217–236. <https://doi.org/10.1007/bf00372820>.
- ZWICKY, Arnold M. & Jerrold M. SADOCK (1975). Ambiguity tests and how to fail them. In *Syntax and semantics* (edited by John Kimball), volume 4, 1–36. New York: Academic Press.