

Distributivity: Debates, advances, questions

Introductory course at ESSLLI at the University of Riga in Latvia

August 5-9, 2019

Lelia Glass

School of Modern Languages at Georgia Institute of Technology (Atlanta, Georgia, USA)

glasslelia@gmail.com

Thanks to the ESSLLI organizers (Jurgis Škilters, Louse McNally, the team at the University of Latvia in Riga, and the organizing committee), as well as to the students and instructors for a very stimulating environment. Thanks to the School of Modern Languages at Georgia Tech for travel funding. I am also grateful to all the people and institutions thanked in my dissertation, Glass 2018b, which this ESSLLI course draws on, particularly my dissertation committee (Beth Levin, Christopher Potts, Cleo Condoravdi, Daniel Lassiter) as well as Lucas Champollion and Hanna de Vries. I am very pleased to do linguistics in a country where the current Prime Minister himself has a Ph.D. in linguistics!

Contents

1	Key facts, terms, and debates	5
1.1	Introduction	5
1.1.1	What is distributivity?	5
1.1.2	Why is distributivity interesting?	7
1.1.3	Goals of this course	7
1.2	Distributivity in more detail	8
1.2.1	The subject	8
1.2.2	The object	8
1.2.3	What's possible vs. what's preferred	11
1.2.4	'Atomic' vs. 'non-atomic' distributivity	12
1.3	What is distributivity contrasted with?	14
1.4	Wrap-up	16
1.4.1	Terminology	16
1.4.2	⚡Discussion	18
2	Semantic analyses	19
2.1	Introduction	19
2.2	Approaches to the semantics of distributivity	20
2.2.1	One source of distributivity: Distributivity as plurality	20
2.2.2	One source of distributivity: Meaning postulates	21
2.2.3	Two sources of distributivity: Meaning postulates and a <i>D</i> operator	23
2.2.4	One source of distributivity: Distributivity as one setting of a 'cover' of the subject	25
2.3	Wrap-up	27
2.3.1	Terminology	27
2.3.2	⚡Discussion	29
3	Which verb phrases go which ways and why?	31
3.1	Motivation	31

3.2	Predictions	32
3.3	Distributivity Ratings Dataset	33
3.4	Testing the predictions	36
3.4.1	Statistical analysis	36
3.4.2	Transitive/intransitive asymmetry	38
3.4.3	Body-mind generalization	40
3.4.4	Multilateral generalization	40
3.4.5	Causative generalization	41
3.4.6	Discussion	42
3.5	Wrap-up	43
3.5.1	Terminology	43
3.5.2	↪Discussion	44
4	(Stubborn) distributivity among adjectives	47
4.1	Introduction	47
4.2	Measurement theory	49
4.3	Explaining the distributivity potential of adjectives	50
4.3.1	Related work	54
4.4	Wrap-up	55
4.4.1	Terminology	55
4.4.2	↪Discussion	56
5	Review and research brainstorm	59
5.1	Introduction	59
5.2	Terminology	60
5.3	Big questions ★	64
5.4	Discussion	71

Day 1

Key facts, terms, and debates

1.1 Introduction

1.1.1 What is distributivity?

- Goal of semantics/pragmatics: Understand inferences that people draw from (uses of) sentences
- When a sentence has a plural subject, we draw inferences about how each member of the subject participates in the predicate of the sentence¹
 - **Distributive:** True of each individual member of a plural subject
 - **Nondistributive:** True of a plural subject as a whole but *not* each individual member

(1) Alice and Bob smiled. **distributive**

- a. ✓**Distributive:** Alice and Bob *each* smiled.
- b. ✗**Nondistributive:** Alice and Bob smiled jointly without each individually doing so.

– ⚡ Can you brainstorm 3 more predicates that behave like *smile*?

(2) Alice and Bob met. **nondistributive**

- a. ✗**Distributive:** Alice and Bob *each* met.
- b. ✓**Nondistributive:** Alice and Bob met jointly without each individually doing so.

– ⚡ Can you brainstorm 3 more predicates that behave like *meet*?

¹Key references include Link 1983, Dowty 1987, Roberts 1987, Landman 1989a, Lasersohn 1995, Schwarzschild 1996, Moltmann 1997, 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.

- (3) Alice and Bob opened the window. **both ways**
- a. ✓**Distributive:** Alice and Bob *each* opened the window.
 - b. ✓**Nondistributive:** Opened the window jointly without each individually doing so.

– ⚡ Can you brainstorm 3 more predicates that behave like *open the window*?

- Focusing mainly on ‘covert’ distributivity (Champollion 2010): inferences drawn in the absence of any disambiguating lexical item such as *each* or *together*
 - vs. ‘overt’ distributivity’: marked with e.g. *each*, *together*
 - ... although the topics are clearly related
- Already, many questions ...
 - Where do these inferences come from? (Silent) logical operators? Facts about the world (people have their own faces so must *smile* individually)? ★
 - * Are these linguistic inferences or non-linguistic inferences? ★
 - (I) **Compositional semantics question:** How to represent these inferences? ★
 - * Do the two ways of understanding (3) arise from ambiguity (of what sort)? From under-specification?
 - Throughout, I use the word *understanding* to remain neutral about that question
 - * If (3) is ambiguous, why do (1) and (2) seem unambiguous? ★
 - (II) **Lexical semantics question:** Which predicates are understood in which way(s), and why? ★
 - * Why do (1)–(3) behave as they do? Which other predicates should act like *smile*, *meet*, or *open the window*, and why? ★
 - * Presumably grounded in what we know about how these events unfold in the world (people have their own faces, so can only *smile* individually); but how can we predict the behavior of a new predicate? ★
 - How do the answers to (I) and (II) interact? ★
 - When a predicate can be understood in multiple ways (3), which understanding is preferred, to what extent, and why? ★

1.1.2 Why is distributivity interesting?

- Linguistic vs. non-linguistic knowledge
- Compositional vs. lexical semantics
- Ambiguity vs. underspecification
- What's available at all vs. what's most accessible/preferred
- Relation between distributivity and plurality (what's marked/unmarked and why?)
- Semantics of plurality in nouns vs. verb phrases vs. adjectives (and the nature of these lexical categories)
- Formalizing a few exemplars vs. predicting the behavior of further predicates

1.1.3 Goals of this course

- After this course, you should be able to ...
 - Use and understand key terminology
 - Critically read papers on distributivity and connect them to larger questions/themes
 - Explain several (formal, experimental) approaches to distributivity, and what they do/do not explain
 - Collect and interpret rich data on distributivity in a language with which you are familiar
 - Articulate some current debates and research questions in this area
 - Identify actionable steps towards asking, answering, and contextualizing your own question
- **Schedule:**
 - Day 1: Introduction to data and terminology
 - Day 2: Semantic analyses of distributivity
 - Day 3: Predicting the distributivity potential of verb phrases
 - Day 4: Predicting the distributivity potential of adjectives
 - Day 5: Review, discussion, and open questions
 - Please see <https://www.essllidistributivity.com/> for **recommended** (but not required) **readings**

- Stars ★ mark **big questions** – areas to pursue in your own research
- Lightning bolts ⚡ mark **active learning** – opportunities to synthesize what you learned to make it stick (you learn more by doing than by listening!)

1.2 Distributivity in more detail

1.2.1 The subject

- Many types of plural subjects ...
 - (4) a. Alice and Bob
 - b. The children
 - c. Three children
 - d. Children
 - e. The children and the adults
 - f. The team / class / choir (syntactically singular but notionally plural; see de Vries 2015)
- Definite plurals (*the children*) give rise to **nonmaximality**: *the children smiled* can be true (or true-enough?) if only 90% of the relevant children smiled (Dowty 1987)
 - How to explain nonmaximality? How does it interact with the actual predicate? (*The glasses are clean* suggests that all of them are clean, vs. *the glasses are dirty* could be true if only some are dirty; Yoon 1996 ...) ★
 - * see Malamud 2012, Križ 2016, and Champollion *et al.* to appear
- In contrast, numerals (*three children*) and conjoined names (*Alice and Bob*) do not involve maximality (why mention specific people/numbers if not all of them participated in the event? – Landman 1989b)
- ⚡ Is *The children smiled* **distributive** or **nondistributive** in a situation where only 90% of the children actually smiled? (What's the relation between distributivity and nonmaximality?)

1.2.2 The object

(Non)distributivity of the object argument

- Mostly, researchers study (non)distributivity of a predicate applied to a subject argument – but objects (and other arguments) can be understood distributively and/or nondistributively too (Dowty 1987, Lasersohn 1993, Champollion to appear)

(5) Alice summarized the proposals. (Dowty 1987)

- a. **✓Distributive:** Summarized each proposal.
- b. **(??) Nondistributive:** Summarized all the proposals (together, into one document) without summarizing each one individually.

- ⚡ Can you brainstorm 3 more predicates that behave like *summarize*? (It might be pretty hard ... why?)

(6) Alice read the proposals.

- a. **✓Distributive:** Read each proposal.
- b. **✗Nondistributive:** Read all the proposals without reading each one individually.

- ⚡ Can you brainstorm 3 more predicates that behave like *read*?

(7) Alice combined the proposals.

- a. **(??) Distributive:** Combined each proposal.
- b. **✓Nondistributive:** Combined all the proposals (together) without combining each one individually.

- ⚡ Can you brainstorm 3 more predicates that behave like *combine*?

Effect of the object on a predicate's distributivity potential (on its subject argument)

- The object of a transitive verb also shapes the distributivity potential of that VP (on its subject argument) ...

- Singular vs. (numeral) plural

(8) Alice and Bob saw a photo.

- a. **✓Distributive:** Each saw a photo.
- b. **✗Nondistributive:** Saw a photo jointly but not individually.

(9) Alice and Bob saw two photos.

- a. ✓**Distributive**: Each saw two photos.
- b. ✓**Nondistributive**: Saw two photos jointly but not individually.

* Perhaps each saw one photo, adding up to two photos between them? (‘Cumulative’; Scha 1981, Krifka 1992)

– World knowledge about the object’s referent

(10) Alice and Bob opened {a soda / a vault / an eye}.

- a. ✓**Distributive**: Each opened {a soda / a vault / an eye}.
- b. (??) **Nondistributive**: Opened {a soda / a vault / an eye} jointly but not individually.

– Definite vs. indefinite – which further interacts with whether the action described by the verb can vs. cannot be repeated on the same object

* which also interacts with **covariation** of the object – whether an indefinite/numeral is taken to ‘covary’ with each member of the subject (Dotlačil 2010) ...

	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. (<i>same window or different ones</i>) ✓ Nondist : Jointly opened one.
Not repeatable on obj. (<i>break</i>)	A&B broke the window. ✗ Dist : Each broke it. ✓ Nondist : Jointly broke it.	A&B broke a window. ✓ Dist : Each broke one. (<i>different windows</i> : ‘covariation’) ✓ Nondist : Jointly broke one.

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

- ⚡ Looking at Table 1.1, why do you think many researchers prefer to use indefinite objects rather than definite objects in constructing examples about distributivity?
- ⚡ Distributivity often involves covarying indefinites, but not always. Can you give an example of a predicate that is understood **distributively** but with no covarying indefinite?

1.2.3 What's possible vs. what's preferred

- When a VP can be understood both distributively and nondistributively, the nondistributive understanding is strongly preferred

– Brooks & Braine 1996, Frazier et al. 1999, Kaup et al. 2002, Dotlačil 2010, Pagliarini et al. 2012, Syrett & Musolino 2013, Maldonado et al. 2017, Champollion to appear

(11) The children built a sand castle. Dobrovie-Sorin et al. 2016

- Distributive (available but dispreferred):** Each built a (different) sand castle.
- Nondistributive (strongly preferred):** They all jointly built a (single) sand castle.

- **Empirical question:** Is there a preference against covarying indefinites; a preference for nondistributivity in general; or both? ★

(12) Alice and Bob opened the door.

- ✓**Distributive:** Each opened the door.
- ✓**Nondistributive:** Opened the door jointly but not individually.

(13) The boxes are heavy. (see Schwarzschild 2011, Scontras & Goodman 2017)

- ✓**Distributive:** Each box is heavy.
- ✓**Nondistributive:** The boxes are jointly heavy but not individually so.

– ↪ Which understanding of (12) is preferred (to what extent)? What about (13)? (We come back to adjectives on Day 4.)

- **Empirical question:** Is the preference for nondistributivity equally strong across languages? (Why or why not?) ★

– Some authors have claimed that in Mandarin, a sentence like (12) can *only* be nondistributive in the absence of the distributivity marker *dōu* (see Lin 1998, Xiang 2008) ...

- **Empirical question:** To what extent does the preference for nondistributivity depend on the actual predicate? (Which aspects of the predicate matter?) ★

– Presumably the indefinite in *A&B ordered a beer* can covary more easily than in *A&B built a sand castle* ...

- **Theoretical question:** How should the observed preference for nondistributivity be explained? (Dotlačil 2010, Dobrovie-Sorin et al. 2016) ★
 - Should distributivity be somehow more formally marked/complex than nondistributivity? (Are more-complex structures actually dispreferred?)
 - Is it somehow easier to process a single event than multiple events? Is it somehow easier to process a single referent for a noun phrase than multiple covarying referents? (Dotlačil 2010)
- Dalrymple et al. 1998 propose the **Strongest Meaning Hypothesis** (based on reciprocals) – that a sentence is understood to encode the logically strongest meaning consistent with the facts
 - *The children know each other* is taken to mean that every child knows every other child
 - *The plates are stacked on top of each other* doesn't mean that every plate is stacked on top of every other plate (not possible), but rather that every plate except the bottom one is stacked on top of some other plate(s)
 - (revised by Kerem et al. 2009 to the **Maximal Typicality Hypothesis** – that a sentence is understood to encode the most 'typical' meaning consistent with the facts; see also Poortman et al. 2018, Winter 2018)
 - **Theoretical question:** Is a distributive understanding of a sentence logically stronger than a non-distributive one? (Or not?) ★
 - * If so, would we expect distributive understandings to be preferred under the Strongest Meaning Hypothesis? (Then why are distributive understandings actually often not preferred?) ★
 - * Under the Maximal Typicality Hypothesis (that hearers prefer to interpret a sentence to describe the most 'typical' situation), would we say that nondistributive understandings are more preferred because they are more 'typical'? If so, why would that be? ★
- In much of literature, authors focus on the understandings that are **available**, regardless of how strongly they are preferred ...

1.2.4 'Atomic' vs. 'non-atomic' distributivity

- So far, we've focused on 'atomic' distributivity: where the predicate 'distributes' to each individual ('atomic') member of a plural subject (e.g., *smile* distributes down to Alice and to Bob)

- But sometimes predicates can be understood to distribute down to non-atomic sub-groups of a plural subject (Gillon 1987, Lasersohn 1989, Gillon 1990, Lasersohn 1995, Schwarzschild 1996, Champollion 2010, Champollion 2019)

- (14) The activists gathered (in a worldwide climate protest). adapted van der Does 1993
- ✗**Atomic distributive:** Each activist gathered.
 - ✓**Non-atomic distributive:** Separate groups of activists gathered in different cities.
 - ✓**Nondistributive:** The activists all gathered jointly but not individually.
- (15) The men wrote operas. (*the men are Handel, Mozart, Gilbert, and Sullivan; Handel wrote operas alone, Mozart wrote operas alone, G&S only wrote operas as a team; example from Gillon 1987*)
- ✓**Atomic distributive:** Each man wrote operas. [false here, but available]
 - ✓**Non-atomic distributive:** Handel wrote operas, Mozart wrote operas; Gilbert and Sullivan wrote operas jointly but not individually.
 - ✓**Nondistributive:** The men wrote operas jointly but not individually. [false here, but available]
- (16) The students gave a presentation.
- ✓**Atomic distributive:** Each student gave a presentation.
 - (??) **Non-atomic distributive:** Each team of 2–3 students gave a presentation.
 - ✓**Nondistributive:** The students gave a presentation jointly but not individually.
- (17) The shoes cost \$50. (*assume 5 pairs of shoes; Lasersohn 1998*)
- (??) **Atomic distributive:** Each shoe costs \$50.
 - ✓**Non-atomic distributive:** Each pair of shoes costs \$50.
 - ✓**Nondistributive:** All the shoes together cost \$50.
- (18) The Teaching Assistants were paid \$14,000. (*assume 3 TAs: Alice, Bob, Caroline; from Lasersohn 1989*)
- ✓**Atomic distributive:** Each TA was paid \$14k.
 - (??) **Non-atomic distributive:** Alice was paid \$14k; Bob and Caroline together were paid \$14k.
 - ✓**Nondistributive:** All the TAs together were paid \$14k.

- Note that non-atomic understandings are much more easily available when the predicate consists of just a verb (*gathered*; (14)) or a verb with a plural object (*wrote operas*; (15); Champollion to

appear)

- ... and much less available for predicates containing indefinites or numerals – especially (18)
 - Previewing semantic theories of distributivity (Day 2), intransitive verbs and VPs with plural objects are often (though not always) assumed to be **cumulative** (Krifka 1992, Kratzer 2007, Champollion 2010) ...
 - * If $a \oplus b$ *gather*, and $c \oplus d$ *gather*, then $a \oplus b \oplus c \oplus d$ also *gather*
 - * If a *smiles* and b *smiles*, then $a \oplus b$ also *smile*
 - * If Mozart *wrote operas* and Handel *wrote operas*, then $m \oplus h$ also *wrote operas* (because plurals are also cumulative: *operas* and more *operas* together are still *operas*)
 - In contrast, numerals and indefinites are not cumulative: if Alice *earned (exactly) 14k* and Bob *earned (exactly) 14k*, then it does **not** follow that $a \oplus b$ *earned (exactly) 14k*
 - ↪ How can (14)–(15) be handled on the assumption that intransitive verbs and predicates with plural objects are **cumulative**? Why doesn't this explanation extend to (16)–(18)?
- **Empirical question:** To what extent are non-atomic understandings (especially non-cumulative ones like (16)–(18)) available/attested – which VPs, and in which contexts? ★
 - **Theoretical question:** If non-atomic understandings (especially non-cumulative ones like (16)–(18)) are somewhat marginal, then to what extent should they be captured in a semantic analysis of distributivity? ★
 - Gillon 1987, Lasersohn 1995, Schwarzschild 1996, Champollion 2016, Champollion 2019
 - **Theoretical question:** How should non-atomic understandings be represented semantically? ★

1.3 What is distributivity contrasted with?

- So far, I've just contrasted **distributive** understandings with **non-distributive** ones
- But there is a debate about what **distributive** understandings should be contrasted with (Roberts 1987, Link 1998b, Landman 2000)
 - Often contrasted with **collective** – but what does that mean?
 - * 'Negative' definition – collective as non-distributive (see Champollion to appear); associated with Roberts 1987, Link 1998b, Kratzer 2007

- * ‘Positive’ definition – collective understandings evoke/entail positive inferences about joint action/responsibility (see Champollion to appear); associated with Landman 2000, Champollion 2010

- Should (the quantificational notion of) distributivity be contrasted with (elusive) notions such as joint action, coordination, and joint responsibility sometimes associated with **collectivity**? Are these ideas parallel enough to be contrasted? (Verkuyl 1994) ★

- * If so, what happens with inanimate/non-agential subjects (19)?

(19) The games entertained the child.

- * (20) is distributive (if two people go running, they each do so); but still defeasibly suggests that Maria and her husband coordinated their actions ...

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

- * (21) is nondistributive (collective, on some definition); but does not involve coordination (White wrote a book expanding a leaflet by his deceased professor Strunk)

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

- One view: two-way split; distributive vs. nondistributive (‘collective’ just a synonym for nondistributive; Roberts 1987, Link 1998b, Kratzer 2007)
- Another view: three-way split; distributive vs. collective (identified by positive inferences about joint action/responsibility) vs. **cumulative** understandings of predicates (Scha 1981)

(22) 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.

(23) 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.

- Are (22b)–(22c) or (23b)–(23c) distinct ‘readings’ of (22), (23)? Or are they just two ways that a nondistributive understanding of (22), (23) could be true?
 - * Again returning to the theme of ambiguity vs. underspecification ...
 - * 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)
 - * ‘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 1998a: Chapter 7)
- For more discussion: Roberts 1987, Link 1998b, Kratzer 2007, Landman 2000, Champollion 2010, Nouwen 2015, Glass 2018b: Chapter 2
- These choices have consequences for the semantic analysis (Day 2); should (22) be three-ways ambiguous? Two-ways ambiguous? Underspecified? ...
 - **Terminology:** Please note that **cumulative** is used in two distinct but related ways in this literature (see Krifka 1992 ...)

(24) **Cumulative** (a property of predicates; Quine 1960)

- a. A predicate P is cumulative iff $P(a) \wedge P(b) \rightarrow P(a \oplus b)$
- b. **Example:** *Water* is cumulative; if a is water and b is water, then $a \oplus b$ is water

(25) **Cumulative** (a property of understandings/readings of predicates; Scha 1981)

- a. A predicate P is understood cumulatively if it is true of a plural subject as a whole, but not individually of each part; usually involves (numeral) plurals in both the subject and the object interacting in a non-scopal manner (Champollion to appear)
- b. **Example:** *Three kids invited six adults* is understood cumulatively if three kids engaged in inviting, and six adults were invited in all.

1.4 Wrap-up

1.4.1 Terminology

1. **Distributive** – a predicate is understood **distributively** if it is inferred to be individually true of each member of a plural subject

2. **Collective** – depending on one’s view, can indicate the absence of distributivity; or can indicate the presence of purported inferences about joint action/responsibility
3. **Cumulative** (property of predicates) – P is cumulative iff $P(a) \wedge P(b) \rightarrow P(a \oplus b)$
4. **Cumulative** (property of readings/understandings of predicates) – P is understood cumulatively if it’s true of a plural subject as a whole, but not individually of each part, such that each member of the plural subject carried out part of the action described by the predicate, jointly adding up to the whole
5. **Overt distributivity** – distributivity marked with disambiguating lexical items such as *each*, *together*
6. **Covert distributivity** – distributivity without any overt markers
7. **Distributive predicate** – a predicate that is only (mainly) understood distributively, such as *smile*
8. **Collective predicate** – a predicate that is only (mainly) understood nondistributively, such as *meet*
9. **Mixed predicate** – a predicate that can be understood both distributively and nondistributively, such as *open the door*
10. **Atom predicate** (Winter 2002): if *all the children PRED* is equivalent to *every child PRED*, then PRED is an atom predicate (e.g., *smile*); if not, it is a set predicate (*numerous*). (Note that it is not clear how to classify ‘mixed’ predicates such as *open the window* in this system; please see de Vries 2015). Atom predicates are said to be those that apply to semantic ‘atoms’ such as *Alice*.
11. **Set predicate** (Winter 2002): if *all the children PRED* is equivalent to *every child PRED*, then PRED is an atom predicate (e.g., *smile*); if not, it is a set predicate (*all the children are numerous* is not equivalent to *?every child is numerous*). (Note, again, that it is not clear how to classify ‘mixed’ predicates such as *open the window* in this system; please see de Vries 2015). Set predicates are said to be those that apply to ‘sets’ (or plurals) such as *the children* rather than to individual atoms such as *Alice*.
12. **Non-maximality** – when some members of a plural subject are excepted from the predicate, as when *the children smiled* can be true(enough) when only 90% of the children actually smiled
13. **Distributive with covariation** – when an indefinite/numeral inside a predicate is understood to ‘covary’ with each member of a plural subject, as when *Alice and Bob drank a beer* means they each drank a different one

14. **(Non)distributive on the subject argument** – inferences drawn about how the predicate of a sentence applies to its plural subject
15. **(Non)distributive on the object argument** – inferences drawn about how the action described in a sentence applies to a plural object of a transitive verb
16. **Atomic distributivity** – when a predicate distributes down to (is individually true of) each ‘atomic’ member of a plural subject
17. **Non-atomic distributivity** – when a predicate distributes down to (is individually true of) sub-groups of a plural subject, but not individual ‘atoms’; e.g. *the shoes cost \$50*

1.4.2 ⚡ Discussion

- (Please discuss in small groups)
- ⚡ Check off all of the terms 1–17 that you feel confident in. Which ones do you feel less confident in? Why?
- ⚡ What did you know/believe about distributivity prior to today’s lesson? What do you know about it now? How does your new knowledge fit into your broader understanding of semantics?
- ⚡ Link 1998b notes that *each* (and distributivity more generally) has a clear quantificational definition, while *together* (and nondistributivity more generally) is compatible with multiple different situations and not as clearly defined. Why is this so? ★
 - Is there a language where the equivalent of *together* unambiguously conveys collectivity? ★
- ⚡ Please go back to one of the questions from today marked with ★. What data or theoretical concerns could shed light on that question? Why might it be interesting? How could you make progress?

Day 2

Semantic analyses

2.1 Introduction

- **Remember from last time:** When a sentence has a plural subject, we draw inferences about how each member of the subject participates in the predicate of the sentence

- **Distributive:** True of each individual member of a plural subject
- **Nondistributive:** True of a plural subject as a whole but *not* each individual member

(1) Alice and Bob smiled. **distributive**

- a. ✓**Distributive:** Alice and Bob *each* smiled.
- b. ✗**Nondistributive:** Alice and Bob smiled jointly without each individually doing so.

(2) Alice and Bob met. **nondistributive**

- a. ✗**Distributive:** Alice and Bob *each* met.
- b. ✓**Nondistributive:** Alice and Bob met jointly without each individually doing so.

(3) Alice and Bob opened the window. **both ways**

- a. ✓**Distributive:** Alice and Bob *each* opened the window.
- b. ✓**Nondistributive:** Opened the window jointly without each individually doing so.

- **Questions for today:** How is distributivity represented (or not represented) semantically? ★
 - Do the two (distributive, nondistributive) understandings of (3) arise from ambiguity, or underspecification? ★
 - If (3) is ambiguous, why are (1)–(2) seemingly unambiguous? ★

- Are the inferences in (1)–(3) fundamentally **linguistic** (i.e., represented in the semantics of these sentences) or **non-linguistic** (i.e., not necessarily represented in the semantics?) ★
- On each of these answers to the **Compositional Semantics Question** of how distributivity should be represented, how do we address the **Lexical Semantics Question** of which predicates are understood in which way(s) and why? ★

2.2 Approaches to the semantics of distributivity

- **Today:** Review four main approaches to the semantics of distributivity
 - (I) One source of distributivity: Distributivity as plurality (Landman 1989a, Landman 2000)
 - (II) One source of distributivity: Meaning postulates (Scha 1981)
 - (III) Two sources of distributivity: Meaning postulates and a *D* operator (Link, Roberts 1987, Champollion 2010 *et seq.*, others)
 - (IV) One source of distributivity: Distributivity as one setting of a ‘cover’ of the plural subject (Gillon 1987, Schwarzschild 1996, Glass 2018b)
- These approaches all assume that distributivity stems primarily located from the predicate rather than the subject (see Champollion to appear for discussion)
 - vs. a minority view (Bennett 1974, Ouwayda 2014) that it distributivity inferences stem from the subject noun phrase
 - ...making it hard to explain how distributive and nondistributive predicates can be coordinated (*the children smiled, met, and opened the door*; Dowty 1987, Lasersohn 1995)

2.2.1 One source of distributivity: Distributivity as plurality

- Distributivity and plurality are ‘two sides of one and the same coin’ (Landman 1989a: 591), handled in a unified manner (Landman 1989a, Landman 2000)
- Singular count nouns \approx ‘singular’ verbs/VPs; plural count nouns \approx ‘pluralized’ verbs/VPs

$$(4) \quad \llbracket child \rrbracket = \{a, b\}$$

$$(7) \quad \llbracket smile \rrbracket = \{a, b\}$$

$$(5) \quad \llbracket \star child \rrbracket = \{a, b, a \oplus b\}$$

$$(8) \quad \llbracket \star smile \rrbracket = \{a, b, a \oplus b\}$$

$$(6) \quad child(a) \wedge child(b) \leftrightarrow \star child(a \oplus b)$$

$$(9) \quad smile(a) \wedge smile(b) \leftrightarrow \star smile(a \oplus b)$$

- Like count nouns (*child*) verbs/VPs such as *smile* are **singular** – meaning that they apply only to ‘atoms’ such as *Alice*, not pluralities such as $Alice \oplus Bob$
- To combine a verb/VP to a plural subject ($Alice \oplus Bob$), the verb/VP must be simultaneously made **plural** and **distributive** via Link’s pluralizing \star operator (9)
 - ... which ensures that the singular version of the verb/VP is individually true of each member of the plural subject
- *Smile* is lexically restricted to take only pure atoms in its denotation, so that it must be pluralized (and made distributive) with \star to combine with a plural subject:

(10) $\star smile(Alice \oplus Bob)$

- *Meet* is lexically restricted to take only groups in its denotation (which behave like atoms but have parts; Link 1983, Landman 1989a); so it acts singular, applying to a group subject without being pluralized
 - using the \uparrow operator of Link 1983, which maps a plural to a group

(11) $meet(\uparrow(Alice \oplus Bob))$

- *Open the window* can take both atoms and groups in its denotation, making it ambiguous between a plural/distributive reading (12a) and a singular/collective reading (12b)

(12) Alice and Bob opened the window.

- Distributive:** $\star open\ the\ window(a \oplus b)$
- Collective:** $open\ the\ window(\uparrow(a \oplus b))$

- The question ‘Which predicates are understood in which ways?’ is a question of which predicates take atoms and/or groups in their denotations
- ⚡ How does Landman achieve his goal of handling distributivity and plurality as ‘two sides of one and the same coin’?

2.2.2 One source of distributivity: Meaning postulates

- Distributivity comes from **meaning postulates** (Carnap 1952) – restrictions on the models that we entertain, in a model-theoretic framework (Scha 1981)

- meant to represent world knowledge about the event described by the predicate (Champollion 2010: 159)
- (13) is distributive thanks to a meaning postulate ensuring that whenever a group G smiles, every member of G also smiles – reflecting the world knowledge that people can only *smile* individually

(13) Alice and Bob smiled.
 $smile(Alice \oplus Bob)$
Meaning postulate: $smile(G) \rightsquigarrow \forall x[x \in G \rightarrow smile(x)]$
- (14) is nondistributive because there is no meaning postulate making it distributive

(14) Alice and Bob met.
 $meet(Alice \oplus Bob)$
- Meaning postulates acknowledge that the inferences drawn from a plural subject depend fundamentally on world knowledge (Roberts 1987, Dowty 1987, de Vries 2015)
 - Dowty 1987: even nondistributive predicates give rise to certain (distributive) inferences about each member of a plural subject – **distributive subentailments**
- (15) The children gathered in the garden. Dowty 1987
Nondistributive: the children gathered jointly without each individually gathering
Distributive subentailment: *each* child was in the garden at the relevant time
- (16) These 3 positive integers sum to 13. Rich Thomason p.c. to Dowty 1987
Nondistributive: the integers sum to 13 jointly without each individually doing so
Distributive subentailment: one or three of the integers is odd
- The question ‘Which predicates are understood in which ways?’ is a question of which meaning postulates are needed in view of what we know about the world
- But this analysis has problems (and no current author uses it) . . .
 - Meaning postulates cannot be optional, so there is no way to explain predicates that can be understood both distributively and nondistributively (*open the door*; Roberts 1987, Champollion to appear, de Vries 2017)

- (17) Alice and Bob opened the door.

$open(Alice \oplus Bob, \iota y[door(y)])$

Strange meaning postulate: $open(G, y) \rightsquigarrow \forall x[x \in G \rightarrow open(x, y)]$

- Also, there is no way to get operators (numerals, indefinites) to covary with each member of a plural subject; $\exists y[raft(y)]$ should take ‘narrow scope’, but with respect to what? (Winter 2001, de Vries 2017)

- (18) Alice and Bob built a raft.

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

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

2.2.3 Two sources of distributivity: Meaning postulates and a *D* operator

- A synthesis of the meaning postulate analysis (for predicates that are always distributive; §2.2.2) and the \star operator analysis (for predicates that can be understood in multiple ways; §2.2.1)
- (I) Some predicates are distributive purely in view of their lexical semantics/what we know about the events they describe
 - *smile* is distributive (formalized via a meaning postulate) because of what we know about faces
 - **P-distributivity** for de Vries 2015; **lexical distributivity** for Champollion 2010 *et seq.*
- (II) Other predicates (where our world knowledge leaves both distributive and nondistributive understandings available) are **ambiguous** between ...
 - **a distributive reading** – derived using the *D* operator, essentially a silent version of *each* (Link 1991, Roberts 1987, Champollion 2010, Champollion 2016)
 - * which can be formulated with or without the event semantics of Davidson 1967, Parsons 1990; see Champollion to appear, Champollion 2016
 - **a nondistributive/collective reading** – in the absence of *D*, and potentially also the presence of the collective/group-forming operator \uparrow (Link 1983, Landman 2000)
 - perhaps also **a cumulative reading** (Scha 1981), derived either via the assumption that all verbs/thematic roles are cumulative (Krifka 1992, Kratzer 2007, Champollion 2010) or via multiple pluralizing \star operators (Beck & Sauerland 2000)

- *Smile* is distributive thanks to a meaning postulate:

- (19) Alice and Bob smiled.

$smile(Alice \oplus Bob)$

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

- *Meet* is nondistributive because it has no meaning postulate (and perhaps also if one adds the group-forming \uparrow operator sometimes attributed to collective predication; Landman 2000)

- (20) Alice and Bob met.

$meet(Alice \oplus Bob)$

or perhaps, depending on one's assumptions: $meet(\uparrow(Alice \oplus Bob))$

- *Open the window* is distributive with *D*, presumably(?) nondistributive without *D* (or perhaps one might add the collective/group-forming \uparrow operator to (21b) too?)

- (21) 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

$open(Alice \oplus Bob, \iota y[window(y)])$

- Covarying indefinites/numerals can take scope underneath *D*

- (22) 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 question ‘Which predicates are understood in which ways?’ is a question of:
 - which meaning postulates to posit?
 - with which predicates is *D* redundant/vacuous, compatible/incompatible?
 - when is \uparrow posited (if at all?) When is it redundant, compatible, incompatible?

2.2.4 One source of distributivity: Distributivity as one setting of a ‘cover’ of the subject

- A cover (Higginbotham 1981, Schwarzschild 1996) is a set of subparts of a plural subject

(23) 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

(24) Covers of {a,b,c}

- a. { {a}, {b}, {c} }
- b. { {a,b,c} }
- c. { {a, b}, {c} }
- d. { {a,b}, {b,c} }
- e. ... (others) ...

- **Applied to a plural subject, a predicate is individually true of each cell of a (contextually supplied) cover of the subject**

– Higginbotham 1981, Gillon 1987, Landman 1996, Schwarzschild 1996, Verkuyl & van der Does 1996

- Originally used primarily for non-atomic distributivity – a cover of (25) can place each pair of shoes in its own cell

(25) The shoes cost \$50.

- Recall from Day 1 that different authors dispute the extent to which non-atomic distributivity exists or should be captured (Gillon 1987, Lasersohn 1989, Gillon 1990, Lasersohn 1995) ★
- Some authors reject the cover analysis, saying that it predicts readings of sentences which they say are not actually available – e.g. *The TAs were paid \$14k* when Alice was paid \$7k, Bob was paid \$7k, and Caroline was paid \$14k
- Others say that those readings are available in a sufficiently rich context; but also that the cover must be one that both the speaker and hearer can coordinate on pragmatically – not every imaginable cover is actually pragmatically available

- * Schwarzschild 1996: speakers/hearers will avoid ‘pathological’ covers such as (24d) (and Malamud 2006 uses game theory to model pragmatic coordination on covers)
- Champollion 2016: if you want to analyze all distributivity using covers, you just have to say that, out of context, the most easily available covers are the ‘endpoints’
 - * A fully distributive cover placing each member of the subject in its own cell
 - * A fully collective cover placing the all in the same cell
 - * Intermediate / non-atomic covers (*the shoes cost \$50*) require more context
- Glass 2018b uses not just non-atomic distributivity or cases where a predicate can be understood in multiple ways (distributively and nondistributively), but for *all* inferences about (non)distributivity
- We know that people have their own faces so can only *smile* individually, so we place each individual in their own cell of the cover:

(26) 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.

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

- We know that people cannot *meet* unilaterally, so we place multiple individuals in the same cell of the cover:

(27) 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

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

- We know that people can open windows individually or jointly, so we entertain a cover where each person opened the window individually, and one where they did so only jointly

(28) Alice and Bob opened the window.

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

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

- The question ‘Which predicates are understood in which ways?’ is a question of which covers to entertain, given what we know about the event described by the predicate

2.3 Wrap-up

2.3.1 Terminology

1. **P-distributivity** – distributivity arising only from the lexical semantics of (or world knowledge about) a particular **Predicate**; e.g. *smile* is said to be P-distributive (Winter 2001, de Vries 2015)
2. **Q-distributivity** – distributivity arising from a **Quantificational** operator in the semantics; e.g. the distributive ‘reading’ (and covarying indefinite) of *Alice and Bob baked a cake* are said to arise from Q-distributivity (Winter 2001, de Vries 2015)
3. **Lexical distributivity** – distributivity arising from the lexical semantics (or world knowledge about) particular lexical item(s); e.g. *smile* is said to be lexically distributive; essentially the same as **P-distributivity** (Champollion 2010 *et seq.*)
4. **Phrasal distributivity** – distributivity of a full multi-word VP containing covarying indefinites or operators; e.g. the distributive (covarying) ‘reading’ of *Alice and Bob baked a cake* is ‘phrasally distributive’; essentially the same as **Q-distributivity**
5. **Distributive subentailment** – an inference which is individually true of each member of a plural subject, but which is not necessarily equivalent to the full predicate of the sentence (Dowty 1987); e.g. *the children gathered in the garden* licenses the ‘distributive subentailment’ that *each* child was in the garden at the relevant time
6. **Meaning postulate** – a stipulated restriction on the models we consider, in a model-theoretic framework. One can use a meaning postulate to ensure that we only consider models where *smile* is distributive (to capture the world knowledge that people have their own faces and so must *smile* individually)
7. **D operator** – essentially, a silent version of *each*, which ensures a distributive understanding/reading of a predicate which could in theory be understood both distributively and nondistributively (Link 1991, Roberts 1987)

8. \uparrow **operator** – used to map sums such as $a \oplus b$ to ‘groups’: $\uparrow (a \oplus b)$ (from Link 1983; see also Landman 2000, Champollion 2010)
9. **Group** – an ‘impure atom’; a composite of multiple ‘atoms’ which itself behaves semantically like an atom of its own; e.g. a and b are atoms, $a \oplus b$ is a sum, $\uparrow (a \oplus b)$ is a group (see Link 1983, Landman 2000, Champollion 2010); used by e.g. Landman 2000 to handle collective ‘readings’ (and associated inferences of joint action/responsibility)
10. \star **operator** – pluralizing operator from Link 1983; $P(a) \wedge P(b) \rightarrow \star P(a \oplus b)$ (if Alice is a rock star, and Bob is a rock star, then Alice and Bob are \star rock stars).
 - [Some technical side notes; feel free to ignore if you just want the big picture]
 - \star is used by Landman 1989a, Landman 2000, de Vries 2015 as a variant of the D operator, yielding a (distributive, plural) reading of a predicate which (in its singular form) applies only to atoms; $smile(a) \wedge smile(b) \rightarrow \star smile(a \oplus b)$.
 - \star is used by Kratzer 2007, Champollion 2010 *et seq.* to indicate that a predicate is assumed to be (inherently) cumulative (in the sense that $P(a) \wedge P(b) \rightarrow P(a \oplus b)$); on the assumption (which not everyone shares!) that verbs are inherently cumulative, $\star smile(a) \wedge \star smile(b) \rightarrow \star smile(a \oplus b)$.
 - but, unlike Landman/de Vries, Champollion/Kratzer do **not** intend for \star to convey distributivity!
 - Champollion/Kratzer assume that **all** verbs are inherently cumulative (an assumption referred to as ‘lexical cumulativity’, which is what they mean when they preface verbs with \star)
 - ...but Champollion/Kratzer certainly do **not** assume that all verbs are inherently distributive! (*meet* is cumulative – if $a \oplus b$ met and $c \oplus d$ met, then $a \oplus b \oplus c \oplus d$ met – but *meet* is **not** distributive: if $a \oplus b$ met, it certainly does not follow that a met and b met).
 - If you want to read carefully on this topic, then it is important to remember that \star always conveys cumulativity – but does not necessarily convey distributivity!
11. **Cover** – a set of subparts of a whole. Used by Schwarzschild 1996 mainly to handle non-atomic distributivity (*the shoes cost \$50*) and sentences that can be understood both distributively and nondistributively (*Alice and Bob opened the window*); used by Glass 2018b to handle **all** distributive and nondistributive understandings of sentences

2.3.2 ↗ Discussion

- (Please discuss in small groups.)
- ↗ Please check off the terms 1–11 that you feel confident in. Which terms do you feel less confident in? Why?
- We've seen four approaches to distributivity:
 - (I) One source of distributivity: Distributivity as plurality (Landman 1989a, Landman 2000)
 - (II) One source of distributivity: Meaning postulates (Scha 1981)
 - (III) Two sources of distributivity: Meaning postulates and a *D* operator (Link, Roberts 1987, Champollion 2010 *et seq.*, others)
 - (IV) One source of distributivity: Distributivity as one setting of a 'cover' of the plural subject (Gillon 1987, Schwarzschild 1996, Glass 2018b, to some extent Champollion 2016)
- ↗ For each approach (I)–(IV), please explain how it would handle the following (understandings of) these sentences:

(29) Caro and Devon laughed.

- a. ✓ **Distributive:** Each laughed.
- b. ✗ **Nondistributive:** Laughed jointly but not individually.

(30) Caro and Devon cooperated.

- a. ✗ **Distributive:** Each cooperated.
- b. ✓ **Nondistributive:** Cooperated jointly but not individually.

(31) Caro and Devon baked a cake.

- a. ✓ **Distributive:** Each baked a (different) cake.
- b. ✓ **Nondistributive:** Baked a cake jointly but not individually.

(32) Caro and Devon saw the photo.

- a. ✓ **Distributive:** Each saw the photo.

b. **✗Nondistributive:** Saw the photo jointly but not individually.

- ⚡ Also, how would each approach (I)–(IV) handle this unusual sentence from Winter & Scha 2015? (What broader lesson can we take from (9))? ★

(33) 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.

- ⚡ Which approach (I)–(IV) do you prefer? Why? ★
- Maldonado et al. 2017 find evidence that distributivity primes distributivity
 - After reading a sentence like *two kids are holding three balloons* in a context where it's (only) true on a distributive understanding (each kid holds three balloons), experimental participants are more likely to interpret an ambiguous/underspecified sentence (*two kids are holding two ice cream cones*) as distributive
 - Distributive understandings primes distributive understandings, while cumulative understandings (*two kids are holding three balloons* where there are 2 kids and 3 balloons total) do not prime cumulative understandings compared to a baseline with no priming – an 'asymmetric' effect of distributivity
 - * Such asymmetric priming effects are often found for less-common or more-marked phenomena
 - Maldonado et al. 2017 argue that this asymmetric priming effect is consistent with analyzing distributivity using a *D* operator – which is marked and can be primed – while treating cumulative readings as unmarked and basic (Krifka 1992, Kratzer 2007)
- ⚡ In your view, to what extent do the findings of Maldonado et al. 2017 help to decide between different semantic analyses (I)–(IV) of distributivity?
- ⚡ Please go back to one of the questions from today marked with ★. What data or theoretical concerns could shed light on that question? Why might it be interesting? How could you make progress?

Day 3

Which verb phrases go which ways and why?

3.1 Motivation

- None of the semantic representations from Day 2 (answers to the **Compositional Semantics Question**) make any predictions about the **Lexical Semantics Question** of which predicates are understood in which ways and why
- Presumably ‘which predicates are understood in which ways?’ depends on world knowledge
 - People have their own faces so can only *smile* individually; an individual person cannot *meet* unilaterally; windows can be opened by both individuals and pluralities thereof
- Many researchers have called for more work on the role of world knowledge in distributivity
 - Dowty’s ‘distributive subentailments’ (Dowty 1987) (1)–(2) stem from world knowledge rather than the logical representation of a sentence; a challenge to explain systematically
 - (1) The children gathered in the garden. Dowty 1987
Nondistributive: the children gathered jointly without each individually gathering
Distributive subentailment: *each* child was in the garden at the relevant time
 - (2) These 3 positive integers sum to 13. Dowty 1987
Nondistributive: the integers sum to 13 jointly without each individually doing so
Distributive subentailment: one or three of the integers is odd

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 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

This [analysis in terms of pragmatically determined covers] leaves us with [...] a pragmatic theory of distributivity, which, **along with other pragmatic phenomena, requires further analysis both in and outside of linguistics.** – Schwarzschild 1996: 101

- For Winter and colleagues, distributivity is one example of **pseudo-quantification** – 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* ‘does not need to be regarded as a truth conditional fact about plurals’ (Winter 2001: 252) but arises from knowledge about sleeping
 - * just as the inference from *the surface is green* to *every part of the surface is green* arises from knowledge about surfaces and greenness rather than from any covert quantifier

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

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

- This is the challenge that I tried to take on in my dissertation (Glass 2018b)

3.2 Predictions

- How to generalize the behavior of *smile*, *meet*, *open the window* beyond a handful of predicates?

- **Body-Mind Generalization:** Because individuals have their own bodies and minds, VPs describing the actions of an individual body or mind (*smile, jump, meditate, see the photo, swallow a pill*) are understood distributively
 - **Multilateral Generalization:** Because individuals cannot carry out inherently multilateral actions alone, VPs describing such actions (*meet*) are understood nondistributively
 - **Causative Generalization:** Because the nature of causation allows that multiple individuals' contributions may be jointly but not individually sufficient to cause a result, causative VPs (describing an event where the subject causes the object to change; *open the window*) allow a nondistributive understanding (in addition, perhaps, to a distributive one)
- Together, these hypothesized patterns also help to explain an otherwise puzzling, striking asymmetry between transitive and intransitive verbs (Link 1983):
 - **Transitive/Intransitive Asymmetry:** Many intransitive verbs (*smile*) are only distributive; while VPs built from many transitive verbs (*open the window*) can be understood nondistributively as well as distributively
 - * Explained indirectly; many intransitive verbs are body-mind verbs (distributive), and many transitive verbs are causative (allowing a nondistributive understanding)

3.3 Distributivity Ratings Dataset

- Glass & Jiang 2017 – summer project with an intern, Nanjiang Jiang, at the Center for the Study of Language & Information at Stanford
- Begin with Levin 1993: list of over 3000 English verbs, categorized by meaning
 - For example: the ‘pour’ verbs, *dribble, drip, pour, slop, slosh, spew, spill, spurt*
- Generate sentences to be rated by human annotators
 - (3) **Stimulus format for VPs built from intransitive and transitive verbs**
 - a. Name1 and Name2 verbed (*Veronika and Ian giggled*).
 - b. Name1 and Name2 verbed an object (*Luke and Olivia wrote a book*).
 - excluding verbs that don’t make sense applied to humans; verbs requiring complements other than a noun
- How to choose objects for transitive verbs?

- Singular count indefinite
 - * because definites add the confounding issue of whether the action described by the verb can be repeated on the same object or not (Day 1)
 - * because plural objects add the potential for a **cumulative** understanding; if Alice opened a door and Bob opened a different door, then perhaps Alice and Bob between them opened *doors* (Day 1)
 - Can't give all verbs the same object! (*open a window* vs. *?eat a window*)
 - Found the nouns most commonly occurring within five words to the right of each verb in the part-of-speech-tagged Spoken section of the Corpus of Contemporary American English (Davies 2008)
 - * If I did this again today, I'd use a dependency parser . . .
 - Then hand-selected the 'best' object from among these
 - * Makes sense as a singular indefinite (*view a photo* over *view a world*)
 - * Fits within the Levin class in which the verb is classified (*snap a twig* over *snap a photo* when *snap* is classified as a change-of-state verb)
 - * Not a body part (*keep a record* over *keep an eye* – to avoid introducing knowledge about bodies; the only exception was the 'verbs involving the body' such as *sprain*, which did get body part objects)
 - * Concrete rather than abstract (*squash a bug* over *squash a hope*)
 - * Not excessively violent or offensive (*persecute a minority* over *persecute a Christian/Jew*)
 - Blend of bottom-up and top-down methods yields objects that are both naturalistically motivated and controlled for confounds
- 2338 unique VPs (1667 transitive verb-object combinations, 671 intransitive verbs)
 - Online participants answered questions of the form (4); five answer choices mapped to a 1–5 Likert scale

(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

- ‘3=not sure’ was the least common response for both questions; most common were ‘4=maybe yes’ and ‘2=maybe no’
 - If I did this again today, I might use unlabeled star ratings (☆☆☆☆☆) so that the mapping to numerals (1–5) for the statistics is clearer ...
 - and/or provide more guidance to participants, because they may have been unsure about how to interpret the questions
 - ⚡(What do you think?)
- (4a) ‘each’ and (4b) ‘together’ probe from two different angles at the same issue: whether the VP is only understood distributively, or whether it can also be understood nondistributively
 - Most VPs describe events that an individual could plausibly undertake individually (*smile*, *open a window*), and thus can be understood distributively when applied to a plural
 - So it is most informative to investigate which VPs have an available nondistributive understanding in addition to a distributive one – to distinguish the *smile* type from the *open the/a window* type
 - Recruited participants using U.S. I.P. addresses on Mechanical Turk; each participant answered 40 random questions of the form (4)
 - Ultimately ran 325 participants total, so that each VP was rated at least 3 times; some VPs were rated more than that (because we didn’t have a way to keep track of how many observations had been recorded for each VP; we kept all the data)
 - ⚡See the full data yourself: <https://osf.io/8953e/>

ParticID	verb	object	‘each’ rating	‘together’ rating	Levin class
Participant1	crack	an egg	2	4	‘bang’
Participant1	cackle	n/a	4	2	‘snap/cackle’
(...)	(...)	(...)	(...)	(...)	(...)
Participant2	steady	a canoe	4	5	‘change-of-state’
Participant2	resent	an intrusion	4	4	‘admire’
(...)	(...)	(...)	(...)	(...)	(...)
Participant3	wheeze	n/a	4	2	‘hiccup’
Participant3	bend	a wire	2	4	‘bend’

Table 3.1: Each participant’s ratings for both the ‘each’ question and the ‘together’ question for each VP they encountered.

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 3.1: Screen shot from the experiment. The optional checkbox allows the participant to indicate that they are unfamiliar with the given verb (some of the verbs, such as *pip*, *confabulate*, and *carom*, were quite rare); if a participant checks that box, their responses for that VP are ignored.

3.4 Testing the predictions

3.4.1 Statistical analysis

- Mixed-effects linear regression (Bates *et al.* 2015): predicts the value of a continuous dependent variable on the basis of one or more independent variables
 - Our dependent variable is not strictly continuous (instead it's ordered/categorical: 'definitely no', 'maybe no', 'not sure', 'maybe yes', and 'definitely yes'), but I treat my 1–5 scale as an approximation of an underlying continuous concept
 - * I also ran all the statistics in mixed-effects ordinal regression models (Cumulative Link Mixed Models; `c1mm` in the ordinal package of R; Christensen 2016)
 - * ...strictly more appropriate for the structure of the ratings scale but harder to graph/interpret because the model outputs log odds ratios rather than predicted 1–5 ratings
 - * All the major findings reported in my linear regressions are replicated in these ordinal logistic regressions¹

¹The Table represents the ordinal regression output for a model (5) predicting the 'each' rating (as an ordered categorical

- Mixed-effects structure (random intercepts for ‘participant’ and ‘VP’) ‘factors out’ differences between participants, VPs unrelated to the hypotheses being tested
- All reported results come from two separate mixed-effects linear regressions – one for the ‘each’ question, one for the ‘together’ question – including all of the independent variables hypothesized to predict distributivity potential:

1. whether the verb is transitive or intransitive
2. whether or not the verb describes an action carried out by an individual body or mind

variable) for a given VP, as a function of whether that VP is tagged as transitive or intransitive; a body/mind verb or not; a multilateral verb or not; causative or not. An interaction between (in)transitivity and body/mind was dropped from the model because it was barely significant at the 0.05 level. The model uses random intercepts for participants and VPs.

```
(5) m <- clmm(each_rating ~trans_intrans + bodymind + multilateral + causative + (1 | SubjId)
+ (1|full_pred))
```

independent var.	estimate	std. error	z	p
transitive	-1.09	0.07	-14.85	***
bodymind	0.98	0.08	12.83	***
multilateral	-0.64	0.14	-4.52	***
causative	-0.35	0.06	-6.26	***
threshold	estimate	std. error	z	
1 2	-4.12	0.12	-35.74	
2 3	-2.37	0.11	-21.63	
3 4	-1.92	0.11	-17.71	
4 5	0.46	0.11	4.33	

A positive estimate means that the independent variable is associated with a higher ‘each’ rating, while a negative estimate means that it is associated with a lower one. (The estimates represent log odds ratios.)

In parallel, the next Table represents the ordinal regression output for a model (6) predicting the ‘together’ rating (as an ordered categorical variable) for a given VP, as a function of whether that VP is tagged as transitive or intransitive; a body/mind verb or not; a multilateral verb or not; causative or not. There are random intercepts for participants and VPs. An interaction between (in)transitivity and body/mind was dropped from the model because it was not significant.

```
(6) m <- clmm(together_rating ~trans_intrans + bodymind + multilateral + causative + (1 |
SubjId) + (1|full_pred))
```

independent var.	estimate	std. error	z	p
transitive	0.77	0.07	10.64	***
bodymind	-1.01	0.08	-13.46	***
multilateral	0.61	0.14	4.33	***
causative	0.24	0.06	4.32	***
threshold	estimate	std. error	z	
1 2	-2.03	0.10	-21.21	
2 3	-0.90	0.09	-9.65	
3 4	-0.45	0.09	-4.83	
4 5	2.04	0.10	31.45	

Again, a positive estimate means that the independent variable is associated with a higher ‘together’ rating, while a negative estimate means that it is associated with a lower one. (The estimates represent log odds ratios.)

The take-home message here is that all the major results reported below from mixed-effects linear regressions are also replicated in ordinal logistic regressions.

3. whether or not the verb describes an inherently multilateral action
 4. whether or not the verb is causative
 5. ...and (only retained if it improved the model according to the Akaike Information Criterion) an interaction between ‘transitive / intransitive’ and ‘body-mind’
- I included all of these independent variables at once in order to isolate the effect of each one – important because they overlap

	trans	intrans	body-mind	multi	causative	total
trans	1667 (100%)	0	112 (7%)	0	945 (57%)	1667
intrans	0	671 (100%)	364 (54%)	91 (14%)	0	671
body-mind	112 (24%)	364 (76%)	476 (100%)	0	0	476
multi	0	91 (100%)	0	91 (100%)	0	91
causative	945 (100%)	0	0	0	945 (100%)	945

Table 3.2: Number of VPs in each category, and overlap between the categories. For example, 945 (57%) of the 1667 transitive verbs are causative.

- All data reported below come from two separate linear regression models, one for the ‘each’ question, one for the ‘together’ question

	predicted ‘each’ rating	β	SE	df	t	p
intercept (regular intrans)	4.08	4.08	0.05	1218	77.4	***
transitive	3.46 (=4.08-0.62)	-0.62	0.05	2245	-13.4	***
body-mind	4.39 (=4.08+0.31)	+0.31	0.05	2167	6.30	***
multilateral (all intrans)	3.68 (=4.08-0.40)	-0.40	0.07	2451	-5.5	***
causative (all trans)	3.28 (=4.08-0.62-0.18)	-0.18	0.03	2182	-5.9	***
trans * body-mind	4.01 (=4.08-0.62+0.31+0.24)	+0.24	0.08	2169	3.1	**

Table 3.3: Mixed-effects linear regression for the ‘each’ question, with random intercepts for both participants and VPs, with an interaction between ‘body-mind’ and ‘transitive’. Statistics below come from this model.

3.4.2 Transitive/intransitive asymmetry

- Link 1983: *carry the piano* can be understood both distributively and nondistributively; while ‘intransitive verbs like *die*’ are only distributive (Link 1983: 132)

- (7) **Transitive/Intransitive Asymmetry:** Most intransitive verbs (*smile*) are only distributive, while VPs built from many transitive verbs (*open the window*) can be understood nondistributively as well as distributively.

	predicted ‘together’ rating	β	SE	df	t	p
intercept (regular intrans)	3.23	3.23	0.05	1277	65.4	***
transitive	3.67 (=3.23+0.44)	+0.44	0.04	2234	10.6	***
body-mind	2.62 (=3.23-0.61)	-0.61	0.04	2201	-14.3	***
multilateral (all intrans)	3.53 (=3.23+0.30)	+0.30	0.08	2460	3.8	***
causative (all trans)	3.81 (=3.23+0.44+0.14)	+0.14	0.03	2206	4.4	***
body-mind * trans	(not included)					

Table 3.4: Mixed-effects linear regression for the ‘together’ question, with random intercepts for both participants and VPs, but no interactions. Statistics below come from this model.

- Is this true? And if it is true, then why?
 - Yes, it’s true: intransitive verbs have higher ‘each’ ratings, VPs built from transitive verbs have higher ‘together’ ratings in the Distributivity Ratings Dataset
 - But why? If the distributivity potential of a VP is grounded in the event it describes, then why would it also be related to argument structure?

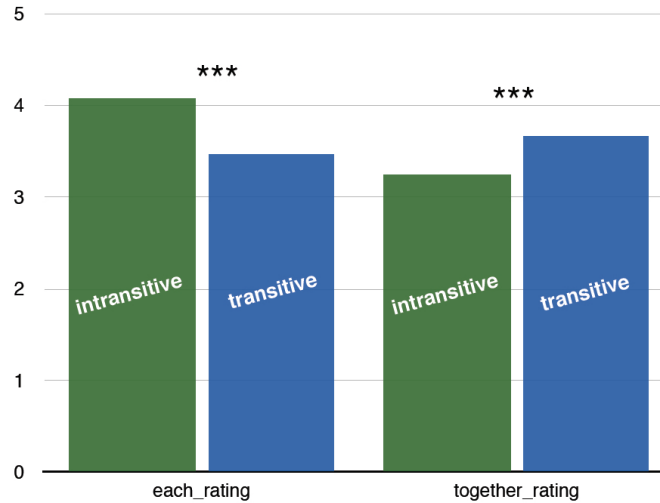


Figure 3.2: Predicted ‘each’ and ‘together’ ratings for transitive and intransitive verbs, from the models described in §3.4.1. VPs built from transitive verbs have systematically lower ‘each’ ratings, and systematically higher ‘together’ ratings, compared to intransitives.

- I claim (Glass 2018b): the relation between argument structure and distributivity is indirect, driven by the types of events that tend to be described by transitive vs. intransitive verbs
 - Some of the more targeted generalizations explored below apply disproportionately to transitives or intransitives, indirectly contributing to this observed asymmetry

3.4.3 Body-mind generalization

- Generalizing the intuitive analysis of *smile* ...
- (8) **Body-Mind Generalization:** Because individuals have their own bodies and minds, VPs describing the actions of an individual body or mind (*smile, jump, meditate, swallow a pill, see a photo, like a book*) are distributive.
- Tagged verbs/VPs in the Distributivity Ratings Dataset that describe bodily/mental actions (using Levin classification as a starting point) – 364 intransitive, 112 transitive; 476 total
 - As predicted, body/mind verbs have higher ‘each’ ratings and lower ‘together’ ratings (compared to non-body/mind verbs)

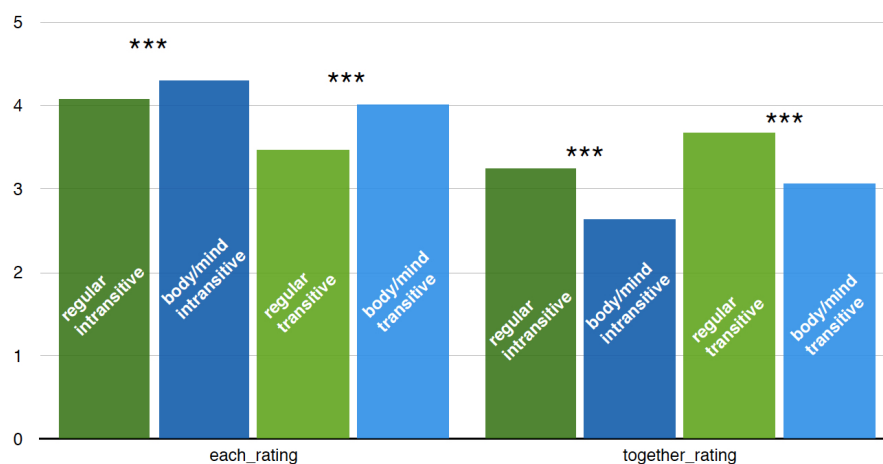


Figure 3.3: Predicted ‘each’ and ‘together’ ratings for body-mind vs. non-body-mind and transitive vs. intransitive verbs, from the models described in §3.4.1. Body-mind intransitives have systematically higher ‘each’ ratings, and systematically lower ‘together’ ratings, than other intransitives. Similarly, body-mind transitives have systematically higher ‘each’ ratings, and systematically lower ‘together’ ratings, than other transitives.

- 76% (364 of 476) of the body/mind VPs are intransitive – indirectly contributing to the observation that many intransitive verbs are distributive

3.4.4 Multilateral generalization

- Generalizing the intuitive analysis of *meet* ...
- (9) **Multilateral Generalization:** Because individuals cannot carry out inherently multilateral actions alone, VPs describing such actions (*meet*) are understood nondistributively.

- Tagged verbs in the Distributivity Ratings Dataset that describe inherently multilateral actions (using Levin classification) – 91 verbs, all intransitive
 - As predicted, multilateral verbs have lower ‘each’ ratings and higher ‘together’ ratings (compared to non-multilateral verbs)

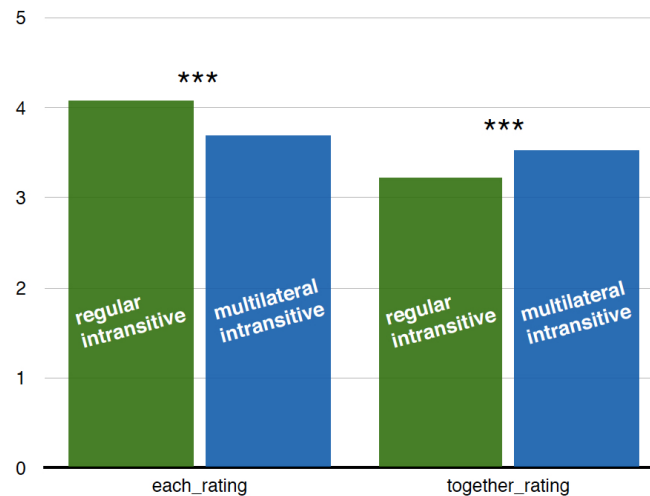


Figure 3.4: Predicted ‘each’ and ‘together’ ratings for multilateral verbs vs. other intransitives, from the models described in §3.4.1. Multilateral verbs have lower ‘each’ ratings and higher ‘together’ ratings than other intransitives.

- These multilateral verbs are all intransitive – **counter to** the observation that many intransitive verbs are distributive

3.4.5 Causative generalization

- Generalizing the intuitive analysis of *open the window . . .* transitive *open* is a **causative** verb (Smith 1970, Dowty 1979) – describing an event in which the subject causes the object to change
 - As a general fact about causation, it is possible for multiple individuals’ actions to jointly bring about a result without each individually doing so
- (10) **Causative Generalization:** Because the nature of **causation** allows that multiple individuals’ contributions may be jointly sufficient but individually insufficient to cause a result, VPs built from **causatives** (*open a window*) are predicted to allow a nondistributive understanding (in addition, perhaps, to a distributive understanding).

- Tagged VPs (built from transitive verbs) in the Distributivity Ratings Dataset that are causative (using Levin classification as a starting point)
 - As predicted, VPs built from causative verbs have lower ‘each’ ratings and higher ‘together’ ratings (compared to non-causatives)

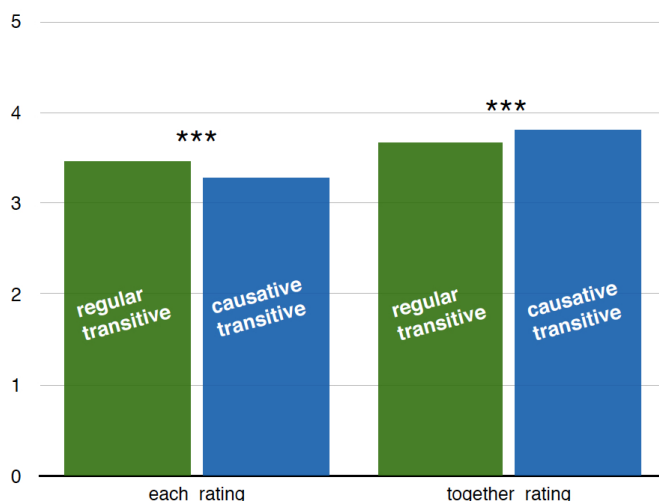


Figure 3.5: Predicted ‘each’ and ‘together’ ratings for causatives vs. other transitive verbs, from the models described in §3.4.1. Causatives (all transitive) have lower ‘each’ ratings and higher ‘together’ ratings than other transitives.

- 57% (945 of 1667) of transitive verbs are causative (and all causatives as defined here are transitive)
 - indirectly contributing to the observation that VPs built from transitive verbs can be understood nondistributively as well as distributively

3.4.6 Discussion

- Still more work to do, more generalizations to be made ...
 - If two individuals are located at a particular place, then they are *each* located at that place (subparts share the location of the whole; Schwarzschild 1996: Chapter 5); thus, spatial predicates also tend to be distributive

- (11) Alice and Bob arrived / departed / entered / exited. **distributive**
- ✓**Distributive:** Alice and Bob *each* arrived / departed / entered / exited.
 - ✗**Nondistributive:** Alice and Bob arrived / departed / entered / exited jointly but not individually.

- Individuals can possess things jointly as well as individually; thus, possession predicates can also be understood nondistributively as well as distributively

(12) Alice and Bob rented a car.

- a. ✓**Distributive:** Alice and Bob *each* rented a car. **both ways**
- b. ✓**Nondistributive:** Alice and Bob rented a car jointly but not individually.

- Two individuals could each affect a different portion of an **incremental** object (an object affected in tandem with the progress of the event of affecting it; Tenny 1992, Krifka 1992), only between them affecting the whole thing; thus, predicates with incremental objects can be understood nondistributively as well as distributively (Glass 2018b)

(13) Alice and Bob ate a pizza.

- a. ✓**Distributive:** Alice and Bob *each* ate a pizza. **both ways**
- b. ✓**Nondistributive:** Alice and Bob ate a pizza jointly (by each eating a different portion of it) but not individually.

- These findings are fairly obvious; but we began with three VPs (*smile, meet, open the window*) and now systematically predict the distributivity of 1512 (476 body-mind VPs, 91 multilateral verbs, and 945 causatives), which is 64% of the 2338 total VPs: progress!

3.5 Wrap-up

3.5.1 Terminology

1. **Pseudo-quantification** – inferences that can be paraphrased using quantificational language without corresponding to any quantifier in the semantic representation (Mador-Haim & Winter 2015)
2. **Transitive/intransitive asymmetry** – the observation (Link 1983, Glass 2017, Glass 2018b) that many intransitive verbs are distributive (*smile*), while VPs built from many transitive verbs can be understood nondistributively as well as distributively (*open the window*)
3. **Body/mind predicate** – a predicate describing the action of an individual body or mind; *smile, swallow a pill*
4. **Multilateral predicate** – a predicate describing an inherently multilateral action; *gather, cooperate*

5. **Causative predicate** – a predicate (transitive verb + object) describing an event in which the subject causes a change to be realized upon the object (Smith 1970)
6. **Incremental object predicate** – a predicate (transitive verb + object) describing an event where the affectedness of the object mirrors the progress of the event of affecting it; *eat the pizza* (Tenny 1992, Krifka 1992)
7. **Mixed-effects model** – a statistical model used for repeated measurements (e.g., from the same item, from the same experimental participant) which accounts for ‘random effects’ (unexplained differences between individual items/participants) in assessing the contribution of ‘main effects’ (independent variables) on the dependent variable
8. **Mixed-effects linear regression** – a statistical model (with a mixed-effects structure, accounting for random effects) which predicts the value of a continuous dependent variable on the basis of one or more independent variables
9. **Mixed-effects ordinal logistic regression** – a statistical model (with a mixed-effects structure, accounting for random effects) which predicts the log odds ratios of each value of an ordered, categorical dependent variable on the basis of one or more independent variables

3.5.2 ↪ Discussion

- (Please discuss in small groups)
 - ↪ Please check off the terms 1–9 that you feel confident in. Which terms do you feel less confident in? Why?
 - In the Distributivity Ratings Dataset, each transitive verb is only tested with one object (*open* is tested with *a window*, not *a soda* or *a vault*)
 - ↪ How could this endeavor be expanded to make/test predictions about the effect of the particular object of the verb? ★
 - ↪ If I (were to) run this Turk study again, how would you suggest that I improve it? Why?
 - ↪ Please explain which understanding(s) are available to (14a)–(14f) and why
- (14) a. Alice and Bob swallowed a pill.

- b. Alice and Bob wrote a book.
 - c. Alice and Bob concocted a plan.
 - d. Alice and Bob emptied a dishwasher.
 - e. Alice and Bob yawned.
 - f. Alice and Bob arrived.

- ⚡ How would your theory represent the fact that causative verbs (*open*) are causative?
 - item Would you use a compositional representation of lexical semantics?
 - Would you state that such verbs just happen to describe events involving causation?
 - Whatever you choose, does your choice make different empirical predictions from other choices?

- ⚡ So far, we've only analyzed sentences carefully constructed from a template. How do the claims made here extend to real data involving sentences with plural subjects (10)? ★
 - (15)
 - a. Several kids pretended to have a picnic at the beach.
 - b. In other U.S. elections, candidates are elected directly by popular vote.
 - c. Consumers chose the most expensive six-pack, because they assume that the greater price indicates greater quality.
 - d. The Stanford doctors tried everything they could to staunch her bleeding.
 - e. We wanted to express our appreciation and thankfulness to you and Dan.
 - f. Teachers practiced some simple experiments that leave a big impact on young learners.

- ⚡ How might it be useful (or not) for an automated natural language understanding system to have information about a predicate's potential for (non)distributivity? ★

- ⚡ In light of today's answers to the **Lexical Semantics Question** of which predicates are understood in which ways, please reconsider from Day 2 the answers to the **Compositional Semantics Question** of how distributivity should be represented semantically.
 - Do today's findings lead you to prefer one semantic analysis over another? Why or why not?

- Kruitwagen *et al.* 2017 finds experimentally that reciprocal predicates such as *hug* do not require symmetric participation from all participants (*Alice and Bob hugged* is still somewhat accepted when Alice hugs Bob but not vice versa)
 - ...but are relatively more accepted when both parties participate actively in the event (see also Winter 2018, Poortman *et al.* 2018)
 - evoking the ‘Maximal Typicality Hypothesis’ (Kerem *et al.* 2009) that sentences are relatively more accepted as true in situations closer to what’s considered ‘typical’ for that verb – *hug* doesn’t strictly entail mutual participation, but that is ‘typical’
 - ⚡How do these findings shed light on the role of lexical semantics and world knowledge in understanding sentences with plural subjects? ★

- ⚡Pasternak 2018 argues that belief verbs (*think, believe*) are not necessarily distributive (11)–(12). Do you take (11)–(12) as evidence against the claim that *think* (as a body/mind verb) is understood distributively? ★

(16) Paul’s cousins think he married a rich New Yorker. Pasternak 2018
where some of his cousins think his spouse is rich, others think the spouse is from NY

(17) Sam’s 6 clients thought she built 6 houses for them. Pasternak 2018
where each client thinks Sam built 1 house for them, and doesn’t know about the others

- ⚡If you could do one large-scale annotation task on Mechanical Turk (similar to the Distributivity Ratings Dataset), what would it be and how would it advance your research?

- ⚡Please go back to one of the questions from today marked with ★. What data or theoretical concerns could shed light on that question? Why might it be interesting? How could you make progress?

Day 4

(Stubborn) distributivity among adjectives

4.1 Introduction

- In addition to verb phrases, predicative adjective phrases can also be understood distributively, nondistributively, or in both ways ...

(1) The boxes are new. **distributive**

- a. ✓**Distributive**: Each box is new.
- b. ✗**Nondistributive**: The boxes are jointly new but not individually so.

– Similarly: *old, light, short, small, full, empty, clean, dirty, fragile* ...

(2) The boxes are connected. **nondistributive**

- a. ✗**Distributive**: Each box is connected.
- b. ✓**Nondistributive**: The boxes are jointly connected but not individually so.

– (Set aside today because they involve a sense of **reciprocity** not shared by the others)

(3) The boxes are heavy. adapted Schwarzschild 1996; **both ways**

- a. ✓**Distributive**: Each box is heavy.
- b. ✓**Nondistributive**: The boxes are jointly heavy but not individually so.

– Similarly: *expensive, beautiful, ugly* ...

- Finally, some adjectives could imaginably be understood **both ways** like (3), but in practice strongly favor a **distributive** understanding (Quine 1960, Schwarzschild 2011): **stubbornly distributive** predicates

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

- Similarly: *big, large, long, wide* ...
- ↪ Can you think of any **stubbornly distributive** verb phrases? Or is this category unique to adjectives?

● **Which adjectives are understood in which ways, and why?**

- Which other adjectives behave like *new, heavy, or tall*?
- Researchers agree that these inferences are grounded in what we know about the property described by the adjective (Dowty 1987, Roberts 1987, Winter 2001 ...) – but otherwise largely left open
- **A recent advance:** Scontras & Goodman 2017 offer a convincing pragmatic explanation for why *heavy* differs from *tall*
 - The imaginable nondistributive understanding of *tall* is difficult to coordinate on because the joint height of boxes depends on transitory spatial arrangement (whereas joint weight is stable) – supported experimentally
 - ↪ Try the experiment here <https://cocolab.stanford.edu/experiments/collective/expt1/expt1.html> (link live as of June 2019)
 - Robot named Cubert handles boxes at a factory – which come out of a dispenser either in a regular stack ('regular') or in a haphazard manner ('random')
 - Each time, Cubert describes the boxes to Dot: *The boxes were heavy / tall / big* ...
 - Web participants rate whether Cubert is describing the boxes as a whole ('collective') or individually ('distributive')
 - For *tall, big*, 'collective' ratings are much higher (nondistributive understandings much more available) when the boxes regularly come out of the machine in a stack
 - Instead of stipulating that *tall, big* are 'stubbornly distributive' while *heavy* is 'complaisantly' nondistributive, Scontras & Goodman derive this pragmatically ...
 - Hearers will not expect a speaker to intend *tall* to be nondistributive, given that the joint height of boxes is transitory; while hearers may expect a speaker to intend *heavy* to be nondistributive, since the joint weight of boxes is consistent

- 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* becomes more available
- *Heavy* is not influenced by the arrangement because the joint weight of boxes does not depend on it
- ⚡ Can you paraphrase in your own words how Scontras & Goodman pragmatically derive the difference between ‘stubbornly distributive’ adjectives such as *tall* and ‘complaisantly nondistributive’ adjectives such as *heavy*?
- What separates predicates like *new* – for which we **cannot even imagine** a nondistributive understanding – from those like *heavy* and *tall*?

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

4.2 Measurement theory

- I propose (Glass 2018a, Glass 2018b) that the answer can be stated using **measurement theory**
 - (Stevens 1946, Suppes & Zinnes 1962, Krantz et al. 1971, Roberts 1979, Krifka 1989, Schwarzschild 2002, Schwarzschild 2006, Sassoon 2007, Sassoon 2010, Lassiter 2011: Chapter 2, Solt 2015 ...)
 - A system used to capture the properties of various sorts of measurements (height, weight, temperature, likelihood, ...)
- Rather than taking (quantitative) numbers/degrees as foundational to measurement, **measurement theory** begins with the (qualitative) notion of relative ordering
- Why not take numbers as basic? Because numbers support relations (e.g. ratio comparisons) that do not carry over to certain qualitative orderings (?*twice as hot*)
- Can construct scales using only as much structure from the natural numbers as appropriate
- Particularly important for **distributivity**, can be used to represent how $\mu(a)$ and $\mu(b)$ relate to $\mu(a \circ b)$...

- where μ maps entities to natural numbers in a way that preserves their relative ordering with respect to one another with respect to the property being measured ...
- ... and $a \circ b$ is the concatenation of a and b – a composite object (if a and b don't overlap, equivalent to Link 1983's 'join' $a \oplus b$; see Lassiter 2011)
- Can define different types of scales based on behavior with respect to concatenation

(5) **Some types of scales**

adapted / abridged Lassiter 2011: 45

- a. **Additive:** $\mu(a \circ b) = \mu(a) + \mu(b)$
Example: weight
- b. **Positive (a.k.a. monotonic; of which additive is a special case):** $\mu(a \circ b) \geq \mu(a)$,
 $\mu(a \circ b) \geq \mu(b)$
Example: loudness, cost
- c. **Intermediate:** If $a \succeq b$, then $\mu(a) \geq \mu(a \circ b) \geq \mu(b)$
Example: temperature (of non-chemically-reacting substances, measured by a thermometer; see Koptjevskaja-Tamm & Rakhilina 2006, Koptjevskaja-Tamm 2011)
- d. **Atom Only:** \succeq_P contains no concatenations; i.e. $a \succeq_P b$ implies that a, b are atoms.
Example: sick, which only makes sense applied to individuals

4.3 Explaining the distributivity potential of adjectives

- Focus on **gradable adjectives** (Bartsch & Vennemann 1972, Cresswell 1976, Klein 1980, Kennedy 1999, Kennedy 2007)
 - Map an entity to its degree d along the scale associated with the adjective
 - *Heavy* is true of *the box* (6) if its weight exceeds contextual standard θ
- (6) $\llbracket \text{The box is heavy} \rrbracket = 1$ iff $\text{heavy}(\text{the box}) \geq \theta_{\text{heavy}}$
- These scales can be characterized using measurement theory – which in turn can explain distributivity potential ...

(7) Box A & Box B are heavy.

(can be understood both ways) (= (3))

- a. ✓**Distributive**: Box A is heavy, Box B is heavy.
- b. ✓**Nondistr.**: Box A & Box B are jointly heavy, individually light.

- **Distributive**: Each box exceeds θ_{heavy}
- **Nondistr.**: A & B together exceed θ_{heavy} , but each individually falls short of it
 - **Plausible because** weight is **additive**; so $a \oplus b$ together surpasses a, b individually

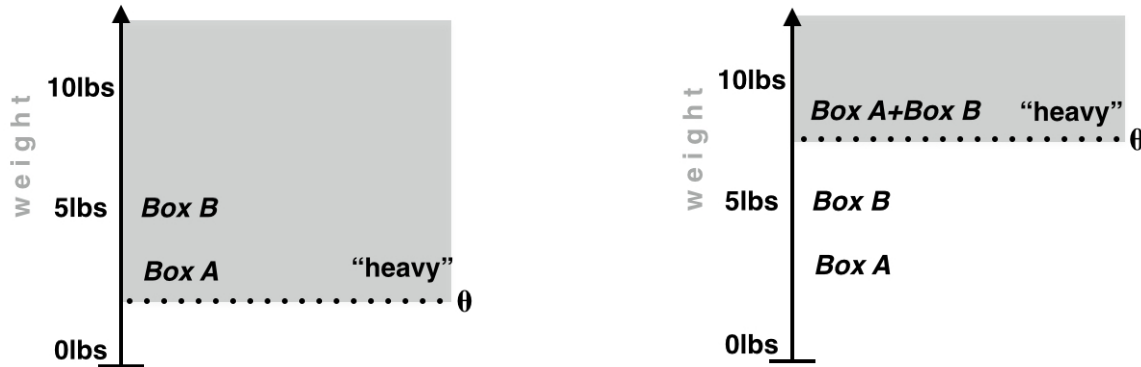


Figure 4.1: Distributive and nondistributive understandings of *heavy*.

- (8) 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 behavior of measurement w.r.t. concatenation, this ordering may or may not be possible
- Other **increasing-direction dimensional adjectives** (*tall, large, expensive*) are also **additive**; thus behave like *heavy* in having an imaginable (though perhaps not pragmatically available) nondistributive understanding
 - As for why *heavy* differs from *tall* – echo Scontras & Goodman 2017 (joint weight is more stable/easier to coordinate on that joint height) ...
 - ...but now we also understand what **unifies** *heavy* and *tall* and **separates** them from *new* – why they both have an imaginable nondistributive understanding
- **Decreasing-direction dimensional adjectives** (*light, short*)

- (9) Box A & Box B are light.

(distributive)

- a. **✓Distributive:** Box A is light, Box B is light.
- b. **✗Nondistr.:** Box A & Box B are jointly light, individually heavy.

- *Light* measures weight in a **decreasing** direction; so a and b are each individually *lighter* than $a \oplus b$ together (Fig. 4.2)
- Cannot set θ_{light} so that $a \oplus b$ exceeds it while a, b individually fall short
- Explains why *light, short* differ from their antonyms in **only** being understood distributively (also *cheap, quiet*)

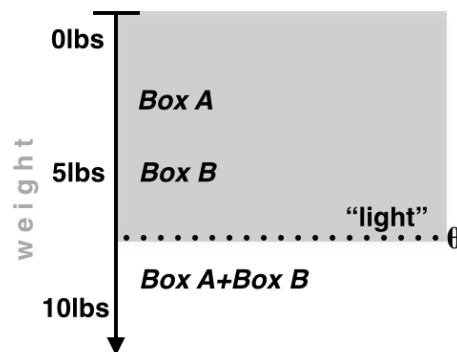


Figure 4.2: *Light* is true of things lighter than the contextual standard θ (here, 7lbs).

• **Adjectives with scales intermediate with respect to concatenation** (*warm*)

(10) Soup A & Soup B are warm. (distributive)

- a. **✓Distributive:** Soup A is warm, Soup B is warm.
- b. **✗Nondistr.:** Soup A & Soup B are jointly warm, individually cool.

- If $\mu(a \oplus b)$ falls in between $\mu(a)$ and $\mu(b)$ (**intermediate**), can't be nondistributive
- No way to set θ so that $a \oplus b$ exceeds it while a, b each fall short
- Similarly: *clean, dirty, old, new* ...

• **Adjectives with scales irregular with respect to concatenation** (*beautiful, ugly*)

(11) Box A & Box B are beautiful. (both ways)

- a. **✓Distributive:** Box A is beautiful, Box B is beautiful.
- b. **✓Nondistr.:** Box A & Box B are jointly beautiful, individually ugly.

- Subjective (Lasersohn 2005); no system to what's *beautiful*; irregular w.r.t. concatenation

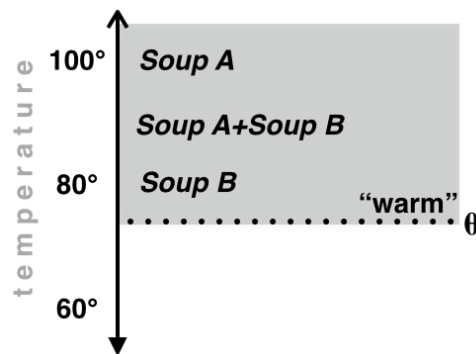


Figure 4.3: Soup A and Soup B are no *warm*er together than individually.

- (Explains why the antonyms *beautiful/ugly* pattern together)
- Similarly: *delicious, disgusting, good, bad* ...
- ‘**Atom-only**’ adjectives (*sick, alive, jealous*)
 - Most adjectives discussed so far have described properties that can be instantiated by pluralities as well as individuals
 - * Boxes can have height/weight/beauty/temperature together as well as individually
 - But (Lassiter 2011: Ch. 2) there are also properties that can only be instantiated by individuals (*sick, alive, worried, religious, American* ...)
 - ... given that individuals have their own bodies/mental states/nationalities
- (12) Alice and Bob are {*sick, alive, American* ...}
 - a. ✓**Distributive**: Each {*sick, alive, American* ...}.
 - b. ✗**Nondistr.**: Jointly {*sick, alive, American* ...}, but not individually so.
- For an adjective to be understood nondistributively, $\mu(a \oplus b)$ must be able to exceed $\mu(a)$, $\mu(b)$ along the scale associated with the adjective – **not possible** if $\mu(a \oplus b)$ undefined
- For some adjectives (*new, full, empty*), not clear if they’re **atom-only** or just **intermediate** w.r.t. concatenation – but same result either way; only **distributive**
- **Predictive, explanatory** answer to ‘Which adjectives are understood in which ways, and why?’

Distributive*Box A & Box B are new (light, short, full, empty).***✓Dist.:** Each new**✗Nondist:** Jointly new**because** $a \oplus b$ can't exceed a, b on *new* scale**Both ways***Box A & Box B are heavy (expensive, beautiful, ugly).***✓Dist.:** Each heavy**✓Nondist:** Jointly heavy**because** $a \oplus b$ can exceed a, b on *heavy* scale

(pragmatically available because joint weight is stable; S&G 2017)

Stubbornly distributive*Box A & Box B are tall (big, large, long, wide).***✓Dist.:** Each tall**(??) Nondist:** Jointly tall**because** $a \oplus b$ can exceed a, b on *tall* scale

(pragmatically unavailable because joint height is unstable; S&G 2017)

4.3.1 Related work

- Some recent papers observe intriguing data on the relation between measurement theory and distributivity – but their analyses to some extent may be seen as restating the data; is it possible to find a more predictive explanation? ★

- Law 2019, building on unpublished work by Linmin Zhang, observes that binomial *each* (*three miles each*) is only available when the noun phrase increases monotonically in relation to the distributive dependency – why? ★

- (13) a. **✓Monotonic NP:** The athletes ran **three miles each**.
 b. **✗Nonmonotonic NP:** The athletes ran **three miles-per-hour each**.

- Schwarzschild 2002, Schwarzschild 2006 observes that while *heavy* can be both distributive and nondistributive in predicative position, it is only distributive in attributive position (the ‘attributive distributive’ generalization) – why? ★

- * Schwarzschild says the dimension described by an attributive adjective *cannot be monotonic* (the joint weight of boxes increases monotonically with the number of boxes, which is not allowed; the weight of an individual box does not increase with the number of boxes, which is allowed)
- * (Interesting that binomial *each* seems to require monotonicity while attributive adjectives seem to require *nonmonotonicity* ... – why? ★)

- (14) The boxes are heavy.

predicative, both ways

- a. **✓Distributive:** Each box is heavy.

- b. **✓Nondistributive:** The boxes are jointly but not individually heavy.
- (15) The heavy boxes (are in the car.) **attributive, distributive**
- a. **✓Distributive:** Each box is heavy.
- b. **✗Nondistributive:** The boxes are jointly but not individually heavy.
- McKinney-Bock & Pancheva 2019 find that *some* attributive adjectives can actually be monotonic (the noisiness of children increases with each child) and nondistributive (children can be jointly but not individually noisy) – why? ★
- (16) The noisy children (are playing.) **both ways**
- a. **✓Distributive:** Each child is noisy.
- b. **✓Nondistributive:** The children are jointly but not individually noisy.
- Contrary to the ‘attributive distributive’ generalization, and contrary to Schwarzschild’s claim that attributive adjectives must describe nonmonotonic properties
- McKinney-Bock and Pancheva say it is because *noisy* modifies events rather than individuals

4.4 Wrap-up

4.4.1 Terminology

1. **Stubbornly distributive adjective** (term from Schwarzschild 2011) – an adjective which *in theory* has an *imaginable* nondistributive understanding, but *in reality* is only understood distributively; e.g. *the boxes are tall*
2. **Complaisantly collective adjective** (term from Scontras & Goodman 2017) – an adjective whose *imaginable* nondistributive understanding is robustly available; e.g. *the boxes are heavy*
3. **Measurement theory** – a system used to capture the properties of various sorts of measurements, based on the non-quantitative concept of relative ordering, using only as much structure from the natural numbers as is suitable to the property being measured
4. **Monotonic / positive with respect to concatenation** – $\mu(a \oplus b) \geq \mu(a), \mu(a \oplus b) \geq \mu(b)$
5. **Additive with respect to concatenation** – $\mu(a \oplus b) = \mu(a) + \mu(b)$
6. **Predicative adjective** – an adjective that serves as the predicate of a sentence, as in *the boxes are heavy*. Depending on the specific adjective, can be distributive or nondistributive

7. **Attributive adjective** – an adjective that does not serve as the predicate of a sentence, as in *the heavy boxes are in the car*. Usually (but not always) distributive (Schwarzschild’s ‘attributive distributive’ generalization)
8. **Binomial *each*** – *each* at the end of a (numeral) noun phrase (*the athletes ran three miles each*)

4.4.2 ↪ Discussion

- (Please discuss in small groups)
 - ↪ Please check off all of the terms 38–8 that you feel confident in. Which ones do you feel less confident in? Why?
 - ↪ Please explain which understanding(s) are available to (17a)–(17f) and why
- (17)
- a. The boxes are wide.
 - b. The boxes are narrow.
 - c. The boxes are huge.
 - d. The boxes are beautiful.
 - e. The boxes are empty.
 - f. The boxes are lacquered.
- Above (Day 1), we saw that when a verb phrase can be understood both distributively and nondistributively, the distributive understanding is **dispreferred** (hearers want *the children built a sand castle* to be nondistributive, with one sand castle)
 - ↪ Among adjectives, is it also true that distributivity is dispreferred? Why or why not? ★
 - Scontras & Goodman 2017 classify decreasing-direction dimensional adjectives (*light, narrow*) as **stubbornly distributive** – so *light* has an imaginable-but-pragmatically-unavailable nondistributive understanding
 - In contrast, Glass 2018a treats *light* like *new*, claiming that it does not have an imaginable nondistributive understanding at all (there is no way for two boxes to be jointly but not individually *light*)

- Maldonado *et al.* 2018 say that *light* is ambiguous between a distributive reading and a collective one – but its collective reading happens to entail its distributive reading (if two boxes are jointly *light* then they are also each *light*)
- ⚡ How would you classify *light*? Why?
- Syrett 2015 finds experimentally that people reject *The boxes are round* in a situation where ten square boxes are arranged into a (round) circle, but accept it when ten round boxes are arranged into a square
 - ⚡ Is this finding consistent with the pragmatic explanation of stubborn distributivity from Scontras & Goodman 2017? Why or why not?
- ⚡ Yoon 1996 observes that different types of adjectives behave differently with respect to (non)maximality (Dowty 1987): *the glasses are clean* suggests that they are all clean, while *the glasses are dirty* might be used even if only 20% of the glasses are actually dirty. Which other adjectives behave like *clean* or *dirty* and why? ★
- ⚡ We saw above (Day 1) that, in the realm of verb phrases, distributivity is marked and dispreferred (Link 1998b, Dotlačil 2010). Does this generalization extend to adjectives? Why or why not? ★
 - (See Syrett 2015, Scontras & Goodman 2017, Maldonado *et al.* 2018 for relevant experimental data)
- ⚡ Please go back to one of the questions from today marked with ★. What data or theoretical concerns could shed light on that question? Why might it be interesting? How could you make progress?

Day 5

Review and research brainstorm

5.1 Introduction

- When a sentence has a plural subject, we draw inferences about how each member of the subject participates in the predicate of the sentence
 - **Distributive:** True of each individual member of a plural subject
 - **Nondistributive:** True of a plural subject as a whole but *not* each individual member

- (1) Alice and Bob smiled. **distributive**
 - a. ✓**Distributive:** Alice and Bob *each* smiled.
 - b. ✗**Nondistributive:** Alice and Bob smiled jointly without each individually doing so.

- (2) Alice and Bob met. **nondistributive**
 - a. ✗**Distributive:** Alice and Bob *each* met.
 - b. ✓**Nondistributive:** Alice and Bob met jointly without each individually doing so.

- (3) Alice and Bob opened the window. **both ways**
 - a. ✓**Distributive:** Alice and Bob *each* opened the window.
 - b. ✓**Nondistributive:** Opened the window jointly without each individually doing so.

Why is distributivity interesting?

- Linguistic vs. non-linguistic knowledge
- Compositional vs. lexical semantics
- Ambiguity vs. underspecification

- What's available at all vs. what's most accessible/preferred
- Relation between distributivity and plurality (what's marked/unmarked and why?)
- Semantics of plurality in nouns vs. verb phrases vs. adjectives (and the nature of these lexical categories)
- Formalizing a few exemplars vs. predicting the behavior of further predicates
- ⚡ Which topics in this class have evoked these themes for you?

5.2 Terminology

1. **Distributive** – a predicate is understood **distributively** if it is inferred to be individually true of each member of a plural subject
2. **Collective** – depending on one's view, can indicate the absence of distributivity; or can indicate the presence of purported inferences about joint action/responsibility
3. **Cumulative** (property of predicates) – P is cumulative iff $P(a) \wedge P(b) \rightarrow P(a \oplus b)$
4. **Cumulative** (property of readings/understandings of predicates) – P is understood cumulatively if it's true of a plural subject as a whole, but not individually of each part, such that each member of the plural subject carried out part of the action described by the predicate, jointly adding up to the whole
5. **Overt distributivity** – distributivity marked with disambiguating lexical items such as *each*, *together*
6. **Covert distributivity** – distributivity without any overt markers
7. **Distributive predicate** – a predicate that is only (mainly) understood distributively, such as *smile*
8. **Collective predicate** – a predicate that is only (mainly) understood nondistributively, such as *meet*
9. **Mixed predicate** – a predicate that can be understood both distributively and nondistributively, such as *open the door*
10. **Atom predicate** (Winter 2002): if *all the children PRED* is equivalent to *every child PRED*, then PRED is an atom predicate (e.g., *smile*); if not, it is a set predicate (*numerous*). (Note that it is not clear how to classify 'mixed' predicates such as *open the window* in this system; please see de Vries 2015). Atom predicates are said to be those that apply to semantic 'atoms' such as *Alice*.

11. **Set predicate** (Winter 2002): if *all the children PRED* is equivalent to *every child PRED*, then PRED is an atom predicate (e.g., *smile*); if not, it is a set predicate (*all the children are numerous* is not equivalent to *?every child is numerous*. (Note, again, that it is not clear how to classify ‘mixed’ predicates such as *open the window* in this system; please see de Vries 2015). Set predicates are said to be those that apply to ‘sets’ (or plurals) such as *the children* rather than to individual atoms such as *Alice*.
12. **Non-maximality** – when some members of a plural subject are excepted from the predicate, as when *the children smiled* can be true(enough) when only 90% of the children actually smiled
13. **Distributive with covariation** – when an indefinite/numeral inside a predicate is understood to ‘covary’ with each member of a plural subject, as when *Alice and Bob drank a beer* means they each drank a different one
14. **(Non)distributive on the subject argument** – inferences drawn about how the predicate of a sentence applies to its plural subject
15. **(Non)distributive on the object argument** – inferences drawn about how the action described in a sentence applies to a plural object of a transitive verb
16. **Atomic distributivity** – when a predicate distributes down to (is individually true of) each ‘atomic’ member of a plural subject
17. **Non-atomic distributivity** – when a predicate distributes down to (is individually true of) sub-groups of a plural subject, but not individual ‘atoms’; e.g. *the shoes cost \$50*
18. **P-distributivity** – distributivity arising only from the lexical semantics of (or world knowledge about) a particular Predicate; e.g. *smile* is said to be P-distributive (Winter 2001, de Vries 2015)
19. **Q-distributivity** – distributivity arising from a Quantificational operator in the semantics; e.g. the distributive ‘reading’ (and covarying indefinite) of *Alice and Bob baked a cake* are said to arise from Q-distributivity (Winter 2001, de Vries 2015)
20. **Lexical distributivity** – distributivity arising from the lexical semantics (or world knowledge about) particular lexical item(s); e.g. *smile* is said to be lexically distributive; essentially the same as **P-distributivity** (Champollion 2010 *et seq.*)
21. **Phrasal distributivity** – distributivity of a full multi-word VP containing covarying indefinites or operators; e.g. the distributive (covarying) ‘reading’ of *Alice and Bob baked a cake* is ‘phrasally distributive’; essentially the same as **Q-distributivity**

22. **Distributive subentailment** – an inference which is individually true of each member of a plural subject, but which is not necessarily equivalent to the full predicate of the sentence (Dowty 1987); e.g. *the children gathered in the garden* licenses the ‘distributive subentailment’ that *each* child was in the garden at the relevant time
23. **Meaning postulate** – a stipulated restriction on the models we consider, in a model-theoretic framework. One can use a meaning postulate to ensure that we only consider models where *smile* is distributive (to capture the world knowledge that people have their own faces and so must *smile* individually)
24. **D operator** – essentially, a silent version of *each*, which ensures a distributive understanding/reading of a predicate which could in theory be understood both distributively and nondistributively (Link 1991, Roberts 1987)
25. **\uparrow operator** – used to map sums such as $a \oplus b$ to ‘groups’: $\uparrow (a \oplus b)$ (from Link 1983; see also Landman 2000, Champollion 2010)
26. **Group** – an ‘impure atom’; a composite of multiple ‘atoms’ which itself behaves semantically like an atom of its own; e.g. a and b are atoms, $a \oplus b$ is a sum, $\uparrow (a \oplus b)$ is a group (see Link 1983, Landman 2000, Champollion 2010); used by e.g. Landman 2000 to handle collective ‘readings’ (and associated inferences of joint action/responsibility)
27. **\star operator** – pluralizing operator from Link 1983; $P(a) \wedge P(b) \rightarrow \star P(a \oplus b)$ (if Alice is a rock star, and Bob is a rock star, then Alice and Bob are \star rock stars).
28. **Cover** – a set of subparts of a whole. Used by Schwarzschild 1996 mainly to handle non-atomic distributivity (*the shoes cost \$50*) and sentences that can be understood both distributively and nondistributively (*Alice and Bob opened the window*); used by Glass 2018b to handle **all** distributive and nondistributive understandings of sentences
29. **Pseudo-quantification** – inferences that can be paraphrased using quantificational language without corresponding to any quantifier in the semantic representation (Mador-Haim & Winter 2015)
30. **Transitive/intransitive asymmetry** – the observation (Link 1983, Glass 2017, Glass 2018b) that many intransitive verbs are distributive (*smile*), while VPs built from many transitive verbs can be understood nondistributively as well as distributively (*open the window*)
31. **Body/mind predicate** – a predicate describing the action of an individual body or mind; *smile*, *swallow a pill*

32. **Multilateral predicate** – a predicate describing an inherently multilateral action; *gather, cooperate*
33. **Causative predicate** – a predicate (transitive verb + object) describing an event in which the subject causes a change to be realized upon the object (Smith 1970)
34. **Incremental object predicate** – a predicate (transitive verb + object) describing an event where the affectedness of the object mirrors the progress of the event of affecting it; *eat the pizza* (Tenny 1992, Krifka 1992)
35. **Mixed-effects model** – a statistical model used for repeated measurements (e.g., from the same item, from the same experimental participant) which accounts for ‘random effects’ (unexplained differences between individual items/participants) in assessing the contribution of ‘main effects’ (independent variables) on the dependent variable
36. **Mixed-effects linear regression** – a statistical model (with a mixed-effects structure, accounting for random effects) which predicts the value of a continuous dependent variable on the basis of one or more independent variables
37. **Mixed-effects ordinal logistic regression** – a statistical model (with a mixed-effects structure, accounting for random effects) which predicts the log odds ratios of each value of an ordered, categorical dependent variable on the basis of one or more independent variables
38. **Stubbornly distributive adjective** (term from Schwarzschild 2011) – an adjective which *in theory* has an *imaginable* nondistributive understanding, but *in reality* is only understood distributively; e.g. *the boxes are tall*
39. **Complaisantly collective adjective** (term from Scontras & Goodman 2017) – an adjective whose *imaginable* nondistributive understanding is robustly available; e.g. *the boxes are heavy*
40. **Measurement theory** – a system used to capture the properties of various sorts of measurements, based on the non-quantitative concept of relative ordering, using only as much structure from the natural numbers as is suitable to the property being measured
41. **Monotonic / positive with respect to concatenation** – $\mu(a \oplus b) \geq \mu(a), \mu(a \oplus b) \geq \mu(b)$
42. **Additive with respect to concatenation** – $\mu(a \oplus b) = \mu(a) + \mu(b)$
43. **Predicative adjective** – an adjective that serves as the predicate of a sentence, as in *the boxes are heavy*. Depending on the specific adjective, can be distributive or nondistributive

44. **Attributive adjective** – an adjective that does not serve as the predicate of a sentence, as in *the heavy boxes are in the car*. Usually (but not always) distributive (Schwarzschild’s ‘attributive distributive’ generalization)
45. **Binomial *each*** – *each* at the end of a (numeral) noun phrase (*the athletes ran three miles each*)

5.3 Big questions ★

Day 1

- Where do these inferences come from? (Silent) logical operators? Facts about the world (people have their own faces so must *smile* individually)? ★
 - Are these linguistic inferences or non-linguistic inferences? ★
- (I) **Compositional semantics question:** How to represent these inferences? ★
 - Do the two ways of understanding (3) arise from ambiguity (of what sort)? From underspecification?
 - * Throughout, I use the word *understanding* to remain neutral about that question
 - If (3) is ambiguous, why do (1) and (2) seem unambiguous? ★
- (II) **Lexical semantics question:** Which predicates are understood in which way(s), and why? ★
 - Why do (1)–(3) behave as they do? Which other predicates should act like *smile*, *meet*, or *open the window*, and why? ★
 - Presumably grounded in what we know about how these events unfold in the world (people have their own faces, so can only *smile* individually); but how can we predict the behavior of a new predicate? ★
- How do the answers to (I) and (II) interact? ★
- When a predicate can be understood in multiple ways (3), which understanding is preferred, to what extent, and why? ★
- Definite plurals (*the children*) give rise to **nonmaximality**: *the children smiled* can be true (or true-enough?) if only 90% of the relevant children smiled (Dowty 1987)

- How to explain nonmaximality? How does it interact with the actual predicate? (*The glasses are clean* suggests that all of them are clean, vs. *the glasses are dirty* could be true if only some are dirty; Yoon 1996 ...) ★
- **Empirical question:** Is the preference for nondistributivity equally strong across languages? (Why or why not?) ★
 - Some authors have claimed that in Mandarin, a sentence like (12) can *only* be nondistributive in the absence of the distributivity marker *dōu* (see Lin 1998, Xiang 2008) ...
- **Empirical question:** To what extent does the preference for nondistributivity depend on the actual predicate? (Which aspects of the predicate matter?) ★
 - Presumably the indefinite in *A&B ordered a beer* can covary more easily than in *A&B built a sand castle* ...
- **Theoretical question:** How should the observed preference for nondistributivity be explained? (Dotlačil 2010, Dobrovie-Sorin et al. 2016) ★
 - Should distributivity be somehow more formally marked/complex than nondistributivity? (Are more-complex structures actually dispreferred?)
 - Is it somehow easier to process a single event than multiple events? Is it somehow easier to process a single referent for a noun phrase than multiple covarying referents? (Dotlačil 2010)
- Dalrymple et al. 1998 propose the **Strongest Meaning Hypothesis** (based on reciprocals) – that a sentence is understood to encode the logically strongest meaning consistent with the facts
 - *The children know each other* is taken to mean that every child knows every other child
 - *The plates are stacked on top of each other* doesn't mean that every plate is stacked on top of every other plate (not possible), but rather that every plate except the bottom one is stacked on top of some other plate(s)
 - (revised by Kerem et al. 2009 to the **Maximal Typicality Hypothesis** – that a sentence is understood to encode the most 'typical' meaning consistent with the facts; see also Poortman et al. 2018, Winter 2018)
- **Theoretical question:** Is a distributive understanding of a sentence logically stronger than a non-distributive one? (Or not?) ★

- If so, would we expect distributive understandings to be preferred under the Strongest Meaning Hypothesis? (Then why are distributive understandings actually often not preferred?) ★
 - Under the Maximal Typicality Hypothesis (that hearers prefer to interpret a sentence to describe a more ‘typical’ situation), would we say that nondistributive understandings are more preferred because they are more ‘typical’? If so, why would that be? ★
- Looking at **non-atomic** distributivity ...
- (4) The activists gathered (in a worldwide climate protest). adapted van der Does 1993
 - a. ✗**Atomic distributive:** Each activist gathered.
 - b. ✓**Non-atomic distributive:** Separate groups of activists gathered in different cities.
 - c. ✓**Nondistributive:** The activists all gathered jointly but not individually.
 - (5) The men wrote operas. (*the men are Handel, Mozart, Gilbert, and Sullivan; Handel wrote operas alone, Mozart wrote operas alone, G&S only wrote operas as a team; example from Gillon 1987*)
 - a. ✓**Atomic distributive:** Each man wrote operas. [**false here, but available**]
 - b. ✓**Non-atomic distributive:** Handel wrote operas, Mozart wrote operas; Gilbert and Sullivan wrote operas jointly but not individually.
 - c. ✓**Nondistributive:** The men wrote operas jointly but not individually. [**false here, but available**]
 - (6) The students gave a presentation.
 - a. ✓**Atomic distributive:** Each student gave a presentation.
 - b. (??) **Non-atomic distributive:** Each team of 2–3 students gave a presentation.
 - c. ✓**Nondistributive:** The students gave a presentation jointly but not individually.
 - (7) The shoes cost \$50. (*assume 5 pairs of shoes; Lasersohn 1998*)
 - a. (??) **Atomic distributive:** Each shoe costs \$50.
 - b. ✓**Non-atomic distributive:** Each pair of shoes costs \$50.
 - c. ✓**Nondistributive:** All the shoes together cost \$50.
 - (8) The Teaching Assistants were paid \$14,000. (*assume 3 TAs: Alice, Bob, Caroline; from Lasersohn 1989*)
 - a. ✓**Atomic distributive:** Each TA was paid \$14k.

- b. (??) **Non-atomic distributive:** Alice was paid \$14k; Bob and Caroline together were paid \$14k.
 - c. ✓**Nondistributive:** All the TAs together were paid \$14k.
- **Empirical question:** To what extent are non-atomic understandings (especially non-cumulative ones like (6)–(8)) available/attested – which VPs, and in which contexts? ★
 - **Theoretical question:** If non-atomic understandings (especially non-cumulative ones like (6)–(8)) are somewhat marginal, then to what extent should they be captured in a semantic analysis of distributivity? ★
 - Gillon 1987, Lasersohn 1995, Schwarzschild 1996, Champollion 2016, Champollion 2019
 - **Theoretical question:** How should non-atomic understandings be represented semantically? ★
 - Is there a connection between (the quantificational notion of) distributivity and (elusive) notions such as joint action, coordination, and joint responsibility sometimes associated with **collectivity**? (Verkuyl 1994) ★
 - ⚡Link 1998b notes that *each* (and distributivity more generally) has a clear quantificational definition, while *together* (and nondistributivity more generally) is compatible with multiple different situations and not as clearly defined. Why is this so? ★
 - Is there a language where the equivalent of *together* unambiguously conveys collectivity? ★

Day 2

- How is distributivity represented (or not represented) semantically? ★
 - (I) One source of distributivity: Distributivity as plurality (Landman 1989a, Landman 2000)
 - (II) One source of distributivity: Meaning postulates (Scha 1981)
 - (III) Two sources of distributivity: Meaning postulates and a *D* operator (Link, Roberts 1987, Champollion 2010 *et seq.*, others)
 - (IV) One source of distributivity: Distributivity as one setting of a ‘cover’ of the plural subject (Gillon 1987, Schwarzschild 1996, Glass 2018b)
- Do the two (distributive, nondistributive) understandings of (3) arise from ambiguity, or under-specification? ★
- If (3) is ambiguous, why are (1)–(2) seemingly unambiguous? ★

- Are the inferences in (1)–(3) fundamentally **linguistic** (i.e., represented in the semantics of these sentences) or **non-linguistic** (i.e., not necessarily represented in the semantics)? ★
- On each of these answers to the **Compositional Semantics Question** of how distributivity should be represented, how do we address the **Lexical Semantics Question** of which predicates are understood in which way(s) and why? ★
- Which approach (I)–(IV) do you prefer? Why? ★
- How would each approach (I)–(IV) handle this unusual sentence from Winter & Scha 2015? (What broader lesson can we take from (9)? ★

- (9) 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.

Day 3

- In the Distributivity Ratings Dataset, each transitive verb is only tested with one object (*open* is tested with *a window*, not *a soda* or *a vault*)
 - ↪How could this endeavor be expanded to make/test predictions about the effect of the particular object of the verb? ★
 - ↪So far, we've only analyzed sentences carefully constructed from a template. How do the claims made here extend to real data involving sentences with plural subjects (10)? ★
- (10)
- Several kids pretended to have a picnic at the beach.
 - In other U.S. elections, candidates are elected directly by popular vote.
 - Consumers chose the most expensive six-pack, because they assume that the greater price indicates greater quality.
 - The Stanford doctors tried everything they could to staunch her bleeding.
 - We wanted to express our appreciation and thankfulness to you and Dan.
 - Teachers practiced some simple experiments that leave a big impact on young learners.
- ↪How might it be useful (or not) for an automated natural language understanding system to have information about a predicate's potential for (non)distributivity? ★

- Kruitwagen *et al.* 2017 finds experimentally that reciprocal predicates such as *hug* do not require symmetric participation from all participants (*Alice and Bob hugged* is still somewhat accepted when Alice hugs Bob but not vice versa)
 - ...but are relatively more accepted when both parties participate actively in the event (see also Winter 2018, Poortman *et al.* 2018)
 - evoking the ‘Maximal Typicality Hypothesis’ (Kerem *et al.* 2009) that sentences are relatively more accepted as true in situations closer to what’s considered ‘typical’ for that verb – *hug* doesn’t strictly entail mutual participation, but that is ‘typical’
 - ⚡ How do these findings shed light on the role of lexical semantics and world knowledge in understanding sentences with plural subjects? ★
- ⚡ Pasternak 2018 argues that belief verbs (*think, believe*) are not necessarily distributive (11)–(12). Do you take (11)–(12) as evidence against the claim that *think* (as a body/mind verb) is understood distributively? ★

- (11) Paul’s cousins think he married a rich New Yorker. Pasternak 2018
where some of his cousins think his spouse is rich, others think the spouse is from NY
- (12) Sam’s 6 clients thought she built 6 houses for them. Pasternak 2018
where each client thinks Sam built 1 house for them, and doesn’t know about the others

Day 4

- Some recent papers observe intriguing data on the relation between measurement theory and distributivity – but their analyses to some extent may be seen as restating the data; is it possible to find a more predictive explanation? ★
 - Law 2019, building on unpublished work by Linmin Zhang, observes that binomial *each* (*three miles each*) is only available when the noun phrase increases monotonically in relation to the distributive dependency – why? ★
- (13) a. ✓ **Monotonic NP:** The athletes ran **three miles each**.
 b. ✗ **Nonmonotonic NP:** The athletes ran **three miles-per-hour each**.
- Schwarzschild 2002, Schwarzschild 2006 observes that while *heavy* can be both distributive and nondistributive in predicative position, it is only distributive in attributive position (the ‘attributive distributive’ generalization) – why? ★

- * Schwarzschild says the dimension described by an attributive adjective *cannot be monotonic* (the joint weight of boxes increases monotonically with the number of boxes, which is not allowed; the weight of an individual box does not increase with the number of boxes, which is allowed)
- * (Interesting that binomial *each* seems to require monotonicity while attributive adjectives seem to require *nonmonotonicity* ... – why? ★)

(14) The boxes are heavy. **predicative, both ways**

- a. ✓**Distributive:** Each box is heavy.
- b. ✓**Nondistributive:** The boxes are jointly but not individually heavy.

(15) The heavy boxes (are in the car.) **attributive, distributive**

- a. ✓**Distributive:** Each box is heavy.
- b. ✗**Nondistributive:** The boxes are jointly but not individually heavy.

- McKinney-Bock & Pancheva 2019 find that *some* attributive adjectives can actually be monotonic (the noisiness of children increases with each child) and nondistributive (children can be jointly but not individually noisy) – why? ★

(16) The noisy children (are playing.) **both ways**

- a. ✓**Distributive:** Each child is noisy.
- b. ✓**Nondistributive:** The children are jointly but not individually noisy.

- Contrary to the ‘attributive distributive’ generalization, and contrary to Schwarzschild’s claim that attributive adjectives must describe nonmonotonic properties
- McKinney-Bock and Pancheva say it is because *noisy* modifies events rather than individuals
- Above (Day 1), we saw that when a verb phrase can be understood both distributively and nondistributively, the distributive understanding is **dispreferred** (hearers want *the children built a sand castle* to be nondistributive, with one sand castle)
 - ⚡ Among adjectives, is it also true that distributivity is dispreferred? Why or why not? ★
 - ⚡ Yoon 1996 observes that different types of adjectives behave differently with respect to (non)maximality (Dowty 1987): *the glasses are clean* suggests that they are all clean, while *the glasses are dirty* might be used even if only 20% of the glasses are actually dirty. Which other adjectives behave like *clean* or *dirty* and why? ★

- ↪ We saw above (Day 1) that, in the realm of verb phrases, distributivity is marked and dispreferred (Link 1998b, Dotlačil 2010). Does this generalization extend to adjectives? Why or why not? ★

* (See Syrett 2015, Scontras & Goodman 2017, Maldonado et al. 2018 for relevant experimental data)

5.4 Discussion

- *(Please discuss in small groups).*
- ⚡ Please check off the terminology that you feel comfortable with. Which do you feel less comfortable with? Why?
- ⚡ Please go back to one of the questions from this week marked with ★. What data or theoretical concerns could shed light on that question? Why might it be interesting? How could you make progress?
- ⚡ If your native language is not English: how does distributivity work in your language?
 - Please consider both grammatical markers of ‘overt’ distributivity as well as unmarked ‘covert’ distributivity
 - and please consider diverse predicates including intransitive verbs; transitive verbs with definite, indefinite, singular, and plural objects; and adjectives
 - Do you observe any differences from English?
- ⚡ What did you think you would learn in this ESSLLI course? What did you actually learn? What will you (hopefully) remember in five years?
- Thank you! Wishing great success to all of you, collectively and distributively

Bibliography

- Bartsch, Renate, & Vennemann, Theo. 1972. The grammar of relative adjectives and comparison. Linguistische Berichte, **20**, 19–32.
- Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software, **67**(1), 1–48. <http://doi.org/10.18637/jss.v067.i01>.
- Beck, Sigrid, & Sauerland, Uli. 2000. Cumulation is needed: A reply to Winter (2000). Natural Language Semantics, **8**(4), 349–371. <http://doi.org/10.1023/a:1011240827230>.
- Bennett, Michael. 1974. Some extensions of a Montague fragment of English. Ph.D. thesis, University of California, Los Angeles.
- Brooks, Patricia J., & Braine, Martin D. 1996. What do children know about the universal quantifiers *all* and *each*? Cognition, **60**(3), 235–268. [http://doi.org/10.1016/0010-0277\(96\)00712-3](http://doi.org/10.1016/0010-0277(96)00712-3).
- Carnap, Rudolf. 1952. Meaning postulates. Philosophical studies, **3**(5), 65–73. <http://doi.org/10.1007/bf02350366>.
- 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**. <http://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. <http://doi.org/10.1093/oso/9780198755128.001.0001>.
- Champollion, Lucas. 2019. Distributivity in formal semantics. Annual Review of Linguistics, **5**. <http://doi.org/10.1146/annurev-linguistics-011718-012528>.
- Champollion, Lucas. to appear. Distributivity, collectivity, and cumulativity. In: Matthewson, Lisa, Meier, Cécile, Rullmann, Hotze, & Zimmermann, Thomas Ede (eds), Companion to Semantics. Hoboken: Wiley.
- Champollion, Lucas, Bumford, Dylan, & Henderson, Robert. to appear. Donkeys under discussion. Semantics and Pragmatics.
- Christensen, Rune Haubo B. 2016. Analysis of ordinal data with cumulative link models – estimation with the R-package ordinal. https://rdrr.io/rforge/ordinal/f/inst/doc/clm_intro.pdf.

- Cresswell, Maxwell John. 1976. The semantics of degree. Pages 261–292 of: Partee, Barbara (ed), *Montague grammar*. New York: Academic Press.
- Dalrymple, Mary, Kanazawa, Makoto, Kim, Yookyung, Mchombo, Sam, & Peters, Stanley. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21(2), 159–210. <http://doi.org/10.1023/a:1005330227480>.
- Davidson, Donald. 1967. The logical form of action sentences. Pages 81–95 of: Rescher, Nicholas (ed), *The logic of decision and action*. Pittsburgh, PA: University of Pittsburgh Press.
- 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. <http://doi.org/10.1075/ijcl.14.2.02dav>.
- 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. <http://doi.org/10.1007/s11050-017-9133-z>.
- Dobrovie-Sorin, Carmen, Ellsiepen, Emilia, & Hemforth, Barbara. 2016. Why are distributive readings dispreferred? Pages 83–102 of: Carrilho, Ernestina, Fiéis, Alexandra, Lobo, Maria, & Pereira, Sandra (eds), *Romance languages and linguistic theory 10: Selected papers from ‘Going Romance’ 28*. Amsterdam: John Benjamins. <http://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.
- Dowty, David R. 1987. A note on collective predicates, distributive predicates, and ‘all’. Pages 97–116 of: Marshall, Fred, Miller, Ann, & Zhang, Zheng-sheng (eds), *Eastern States Conference on Linguistics (ESCOL)*, vol. 3. Columbus: The Ohio State University.
- Frazier, Lyn, Pacht, Jeremy M., & Rayner, Keith. 1999. Taking on semantic commitments, II: Collective versus distributive readings. *Cognition*, 70(1), 87–104. [http://doi.org/10.1016/s0010-0277\(99\)00002-5](http://doi.org/10.1016/s0010-0277(99)00002-5).
- Gillon, Brendan S. 1987. The readings of plural noun phrases in English. *Linguistics and Philosophy*, 10(2), 199–219. <http://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. <http://doi.org/10.1007/bf00630751>.
- Glass, Lelia. 2017. Exploring the relation between argument structure and distributivity. Pages 95–119 of: Nee, Julia, Cychosz, Margaret, Hayes, Dmetri, Lau, Tyler, & Ramirez, Emily (eds), *Berkeley Linguistics Society (BLS)*, vol. 43. Berkeley: eScholarship.
- Glass, Lelia. 2018a. Deriving the distributivity potential of adjectives via measurement theory. In:

- Ferrell, Patrick (ed), *Linguistic Society of America (LSA)*, vol. 3. Washington, D.C.: Linguistic Society of America. <http://doi.org/10.3765/plsa.v3i1.4343>.
- Glass, Lelia. 2018b. *Distributivity, lexical semantics, and world knowledge*. Ph.D. thesis, Stanford University. <http://ling.auf.net/lingbuzz/004050>.
- Glass, Lelia, & Jiang, Nanjiang. 2017. *Distributivity ratings dataset*. Available online through the Open Science Framework at <http://doi.org/10.17605/osf.io/8953e>. Stanford University.
- Higginbotham, James. 1981. Reciprocal interpretation. *Journal of Linguistic Research*, **1**(3), 97–117.
- Kaup, Barbara, Kelter, Stephanie, & Habel, Christopher. 2002. Representing referents of plural expressions and resolving plural anaphors. *Language and Cognitive Processes*, **17**(4), 399–467. <http://doi.org/10.1080/01690960143000272>.
- Kennedy, Christopher. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.
- Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, **30**(1), 1–45.
- Kerem, Nir, Friedmann, Naama, & Winter, Yoad. 2009. Typicality effects and the logic of reciprocity. Pages 257–274 of: Cormany, Ed, Ito, Satoshi, & Lutz, David (eds), *Proceedings of Semantics and Linguistic Theory (SALT) 19*. <https://doi.org/10.3765/salt.v19i0.2537>.
- Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, **1**(4), 1–45.
- Koptjevskaja-Tamm, Maria. 2011. ‘It’s boiling hot’: On the structure of the linguistic temperature domain across languages. Pages 393–410 of: Schmid, Sarah Dessì, Detges, Ulrich, Gévaudan, Paul, Mihatsch, Wiltrud, & Waltereit, Richard (eds), *Rahmen des Sprechens: Beiträge zur Valenztheorie, Varietätenlinguistik, Kognitiven und Historischen Semantik (Frame of speech: Contributions to valence theory, variationist linguistics, cognitive and historical semantics)*. Tübingen: Narr.
- Koptjevskaja-Tamm, Maria, & Rakhilina, Ekaterina V. 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. <http://doi.org/10.1524/stuf.2006.59.3.253>.
- Krantz, David H, Luce, R. Duncan, Suppes, Patrick, & Tversky, Amos. 1971. *Foundations of measurement*. San Diego: Academic Press.
- Kratzer, Angelika. 2007. On the plurality of verbs. Pages 269–300 of: Dölling, Johannes, Heyde-Zybatow, Tatjana, & Schäfer, Martin (eds), *Event structures in linguistic form and interpretation*. Berlin: De Gruyter. <http://doi.org/10.1515/9783110925449.269>.
- Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. Pages 75–115 of: Bartsch, Renate, van Benthem, Johan, & van Emde Boas, Peter (eds), *Semantics and contextual expression*. Berlin: De Gruyter.

- Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. Pages 29–54 of: Sag, Ivan A., & Szabolcsi, Anna (eds), Lexical Matters. Stanford, CA: CSLI (Center for the Study of Language and Information) Publications.
- Križ, Manuel. 2016. Homogeneity, non-maximality, and ‘all’. Journal of Semantics, **3**(33), 493–539. <http://doi.org/10.1093/jos/ffv006>.
- Kruitwagen, Imke, Poortman, Eva B., & Winter, Yoad. 2017. Reciprocal verbs as collective predicate concepts. In: Lamont, Andrew, & Tetzloff, Katerina (eds), North East Linguistic Society (NELS), vol. 47. Amherst: Graduate Linguistics Student Association of the University of Massachusetts.
- Landman, Fred. 1989a. Groups, I. Linguistics and Philosophy, **12**(5), 559–605. <http://doi.org/10.1007/bf00627774>.
- Landman, Fred. 1989b. Groups, II. Linguistics and Philosophy, **12**(6), 723–744. <http://doi.org/10.1007/bf00632603>.
- Landman, Fred. 1996. Plurality. Pages 425–457 of: Lappin, Shalom (ed), The handbook of contemporary semantic theory. Oxford: Blackwell Publishing.
- Landman, Fred. 2000. Events and plurality: The Jerusalem lectures. Dordrecht: Kluwer Academic Publishers.
- Lasersohn, Peter. 1989. On the readings of plural noun phrases. Linguistic Inquiry, **20**(1), 130–134.
- Lasersohn, Peter. 1993. Lexical distributivity and implicit arguments. Pages 145–161 of: Lahiri, Utpal, & Wyner, Adam (eds), Semantics and Linguistic Theory (SALT), vol. 3. Ithaca: Cornell University. <http://doi.org/10.3765/salt.v3i0.2751>.
- Lasersohn, Peter. 1995. Plurality, conjunction and events. Dordrecht: Kluwer Academic Publishers. <http://doi.org/10.1007/978-94-015-8581-1>.
- Lasersohn, Peter. 1998. Generalized distributivity operators. Linguistics and Philosophy, **21**(1), 83–93. <http://doi.org/10.1023/a:1005317815339>.
- Lasersohn, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. Linguistics and Philosophy, **28**(6), 643–686. <http://doi.org/10.1007/s10988-005-0596-x>.
- Lasersohn, Peter. 2011. Mass nouns and plurals. Pages 1131–1153 of: Heusinger, Klaus von, Maienborn, Claudia, & Portner, Paul (eds), Semantics: An international handbook of natural language meaning, vol. 2. Berlin: De Gruyter. <http://doi.org/10.1515/9783110255072.1131>.
- Lassiter, Daniel. 2011. Measurement and modality: The scalar basis of modal semantics. Ph.D. thesis, New York University, New York.
- Law, Jess H. K. 2019. Binominal ‘each’ and association with the internal structure of distributivity. Manuscript under review, Rutgers University / U.C. Santa Cruz.
- Levin, Beth. 1993. English verb classes and alternations. Chicago: University of Chicago Press.
- Lin, Jo-Wang. 1998. Distributivity in Chinese and its implications. Natural Language Semantics, **6**(2), 201–243. <http://doi.org/10.1023/a:1008299031574>.

- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. Pages 127–146 of: Bäuerle, Rainer, Schwarze, Christoph, & Stechow, Arnim von (eds), Meaning, Use and Interpretation of Language. Berlin: De Gruyter.
- Link, Godehard. 1991. Plural. In: von Stechow, Arnim, & Wunderlich, Dieter (eds), Semantik: Ein internationales Handbuch der zeitgenössischen Forschung [Semantics: An international handbook of contemporary research]. Berlin: De Gruyter. <http://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? Pages 19–54 of: Hamm, Fritz, & Hinrichs, Erhard (eds), Plurality and quantification. Studies in Linguistics and Philosophy. Dordrecht: Kluwer Academic Publishers. http://doi.org/10.1007/978-94-017-2706-8_2.
- Mador-Haim, Sela, & Winter, Yoad. 2015. Far from obvious: The semantics of locative indefinites. Linguistics and Philosophy, **38**(5), 437–476. <http://doi.org/10.1007/s10988-015-9175-y>.
- Malamud, Sophia. 2012. The meaning of plural definites: A decision-theoretic approach. Semantics and Pragmatics, **5**. <http://doi.org/10.3765/sp.5.3>.
- Malamud, Sophia A. 2006. (Non)maximality and distributivity: A decision theory approach. Pages 120–137 of: Gibson, Masayuki, & Howell, Jonathan (eds), Semantics and Linguistic Theory (SALT), vol. 16. Ithaca: Cornell Linguistics Circle. <http://doi.org/10.3765/salt.v16i0.2943>.
- Maldonado, Mora, Chemla, Emmanuel, & Spector, Benjamin. 2017. Priming plural ambiguities. Journal of Memory and Language, **95**, 89 – 101. <http://doi.org/10.1016/j.jml.2017.02.002>.
- Maldonado, Mora, Chemla, Emmanuel, & Spector, Benjamin. 2018. Revealing abstract semantic mechanisms through priming: the distributive/collective contrast. Cognition.
- McKinney-Bock, Katy, & Pancheva, Roumyana. 2019. Why is attributive ‘heavy’ distributive? In: Altshuler, Daniel, & Rett, Jessica (eds), The semantics of plurals, focus, degrees, and times: Essays in Honor of Roger Schwarzschild. New York: Springer.
- Moltmann, Friederike. 1997. Parts and wholes in semantics. Oxford: Oxford University Press.
- Nouwen, Rick. 2015. Plurality. Pages 267–284 of: Dekker, Paul, & Aloni, Maria (eds), Cambridge handbook of semantics. Cambridge: Cambridge University Press. <http://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.
- Pagliarini, Elena, Fiorin, Gaetano, & Dotlačil, Jakub. 2012. The acquisition of distributivity in pluralities. Pages 387–399 of: Boston University Conference on Language Development (BUCLD), vol. 36. Somerville: Cascadilla Press.

- Parsons, Terence. 1990. Events in the semantics of English: A study in subatomic semantics. Cambridge, MA: MIT Press.
- Pasternak, Robert. 2018. The mereology of attitudes. Ph.D. thesis, Stony Brook University, Stony Brook, NY.
- Poortman, Eva B., Struiksma, Marijn E., Kerem, Nir, Friedmann, Naama, & Winter, Yoad. 2018. Reciprocal expressions and the Maximal Typicality Hypothesis. Glossa, **3**(1). <http://doi.org/10.5334/gjgl.180>.
- Quine, Willard Van Orman. 1960. Word and object. Cambridge: MIT press.
- Roberts, Craige. 1987. Modal subordination, anaphora, and distributivity. Ph.D. thesis, University of Massachusetts, Amherst.
- Roberts, Fred S. 1979. Measurement theory with applications to decisionmaking, utility, and the social sciences. Addison Wesley Publishing Company.
- 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. <http://doi.org/10.1007/s11229-009-9687-5>.
- Scha, Remko. 1981. Distributive, collective and cumulative quantification. Pages 483–512 of: Groenendijk, Jeroen, & Stockhof, Martin (eds), Formal methods in the study of language. Amsterdam: Mathematical Center Tracts. <http://doi.org/10.1515/9783110867602.131>.
- Schwarzschild, Roger. 1996. Pluralities. Dordrecht: Kluwer Academic Publishers.
- Schwarzschild, Roger. 2002. The grammar of measurement. Pages 225–245 of: Jackson, Brendan (ed), Semantics and Linguistic Theory (SALT), vol. 12. Ithaca: Cornell Linguistics Circle. <http://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. <http://doi.org/10.1111/j.1467-9612.2006.00083.x>.
- Schwarzschild, Roger. 2011. Stubborn distributivity, multiparticipant nouns and the count/mass distinction. In: Lima, Suzi, Mullin, Kevin, & Smith, Brian (eds), North East Linguistic Society (NELS), vol. 39. Amherst: Graduate Linguistics Student Association of the University of Massachusetts.
- Scontras, Gregory, & Goodman, Noah. 2017. Resolving uncertainty in plural predication. Cognition, **168**, 294–311. <http://doi.org/10.1016/j.cognition.2017.07.002>.
- Smith, Carlota S. 1970. Jespersen's 'move and change' class and causative verbs in English. Pages 101–109 of: Jazayery, Mohammad A., Polomé, Edgar C., & Winter, Werner (eds), Linguistic and literary studies in honor of Archibald A. Hill, vol. 2: Descriptive Linguistics. The Hague: Mouton. <http://doi.org/10.1515/9783110800432.101>.
- Solt, Stephanie. 2015. Measurement scales in natural language. Language and Linguistics Compass, **9**(1), 14–32. <http://doi.org/10.1111/lnc3.12101>.

- Stevens, Stanley Smith. 1946. On the theory of scales of measurement. *Science*, **103**(2684), 688–680. <http://doi.org/10.1126/science.103.2684.677>.
- Suppes, Patrick, & Zinnes, Joseph L. 1962. *Basic measurement theory*. Stanford: Institute for Mathematical Studies in the Social Sciences.
- Syrett, Kristen. 2015. Mapping properties to individuals in language acquisition. *Pages 398–410 of: Boston University Conference on Language Development (BUCLD)*, vol. 39. Somerville: Cascadia Press.
- Syrett, Kristen, & Musolino, Julien. 2013. Collectivity, distributivity, and the interpretation of plural numerical expressions in child and adult language. *Language Acquisition*, **20**(4), 259–291. <http://doi.org/10.1080/10489223.2013.828060>.
- Tenny, Carol. 1992. The Aspectual Interface Hypothesis. *Pages 1–28 of: Sag, Ivan A., & Szabolcsi, Anna (eds), Lexical Matters*. Stanford, California: Center for the Study of Language and Information.
- van der Does, Jaap. 1993. Sums and Quantifiers. *Linguistics and Philosophy*, 239–282.
- Verkuyl, Henk. 1994. Distributivity and collectivity: A couple at odds. *Pages 49–80 of: Kanazawa, Makoto, & Piñón, Christopher J. (eds), Dynamics, polarity and quantification*. Stanford: CSLI (Center for the Study of Language and Information) Publications.
- Verkuyl, Henk J., & van der Does, Jaap. 1996. The semantics of plural noun phrases. *In: van der Does, Jaap, & van Eijck, Jan (eds), Quantifiers, logic, and language*. Stanford: CSLI (Center for the Study of Language and Information) Publications.
- Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy*, **20**(4), 399–467. <http://doi.org/10.1023/a:1005354323136>.
- Winter, Yoad. 2000. Distributivity and dependency. *Natural Language Semantics*, **8**(1), 27–69. <http://doi.org/10.1023/a:1008313715103>.
- Winter, Yoad. 2001. *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge: MIT Press.
- Winter, Yoad. 2002. Atoms and sets: A characterization of semantic number. *Linguistic Inquiry*, **33**(3), 493–505. <http://doi.org/10.1162/002438902760168581>.
- Winter, Yoad. 2018. Symmetric predicates and the semantics of reciprocal alternations. *Semantics and Pragmatics*, **11**. <http://doi.org/10.3765/sp.11.1>.
- Winter, Yoad, & Scha, Remko. 2015. Plurals. *Pages 77–113 of: Lappin, Shalom, & Fox, Chris (eds), Handbook of contemporary semantic theory*, 2 edn. New York: Wiley. <http://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. <http://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. <http://doi.org/10.1007/bf00372820>.