Advanced Mereology for Linguists

Lecture notes to the 2021 ESSLLI course

Lucas Champollion
New York University

champollion@nyu.edu

Summer 2021 (Lecture notes version: July 8, 2021)

Preface

This course develops a unified theory of cross-categorial similarities involving the count-mass, singular-plural, telic-atelic, and collective-distributive opposition. The contents are drawn from Champollion 2017. Basic familiarity with formal semantics and mereology, as presented in my previous introductory ESSLLI and LSA course "Linguistic applications of mereology" https://ling.auf.net/lingbuzz/002174, or in the review article Champollion & Krifka 2016, will be helpful.

Day 1: Linguistic applications of mereology.   After an overview of the content to be covered, the course starts by a whirlwind tour of mereology, the singular-plural distinction, the count-mass distinction, higher-order properties, extensive and intensive measure functions, the telic-atelic opposition, and aspectual composition. This is a recap intended to refresh participants' memories; the corresponding topics are covered in more introductory materials such as the first half of the lecture notes at https://ling.auf.net/lingbuzz/002174 and the first few chapters of Champollion 2017.

Day 2: Stratified reference.   This lecture introduces stratified reference, which provides a unified perspective on the domains of measurement, aspect, and distributivity. It explains the linguistic relevance of the difference between extensive measure functions like volume and intensive measure functions like temperature, as illustrated by the pseudopartitives *thirty liters of water* vs. **thirty degrees Celsius of water* (Krifka 1998, Schwarzschild 2006). Subsuming these previous accounts, stratified reference correctly predicts the monotonicity constraint: such constructions disallow measure functions that generally return the same value on an entity and on its parts. For example, in order for **thirty degrees Celsius of water* to be acceptable, it would have to describe a water entity whose parts are colder than itself; but there are no such entities. Stratified reference relativizes unboundedness to just one dimension or measure function at a time. This makes it possible to account for examples like *five feet of snow* even though not every part of a five-foot snow layer of snow is less than five feet high.

Day 3: Collectivity, cumulativity and *all*. This lecture develops an account for differences within the class of collective predicates, as exemplified by the contrast between *all the students gathered* and **all the students were numerous* (Dowty 1987, Winter 2001), and for the limited ability of *all* to take part in cumulative readings. Stratified reference is used to formulate meaning postulates that capture the fact that predicates like *gather* give rise to distributive inferences to subgroups, and to formulate the semantics of *all* in terms of a subgroup distributivity requirement.

Day 4: Covert distributivity. Building on Champollion 2016a, this lecture considers how verb phrases such as *build a raft* optionally acquire a distributive interpretation, and reformulates the covert distributivity operators of Link 1983 (the atomic D operator) and Schwarzschild 2006 (the nonatomic Part operator) in terms of the two parameters of stratified reference. By varying the granularity parameter, the difference between atomic and nonatomic views of distributivity is captured and clarified. By varying the dimension parameter, these distributivity operators are extended to the temporal domain and used to explain why indefinites in the syntactic scope of *for*-adverbials tend not to covary with them (*?John found a flea on his dog for a month*, Zucchi & White 2001).

Day 5: Overt distributivity. Building on Champollion 2016b, this lecture explains the crosslinguistic semantic differences between distance-distributive items such as English *each* and German *jeweils* by treating them as overt versions of the atomic distributivity operator D and the nonatomic distributivity operator Part respectively. The proposed analysis in terms of stratified reference explains why *jeweils* can distribute over salient occasions and why this is never possible for *each* (Zimmermann 2002). It also accounts for the fact that distributive determiners can take part in cumulative readings with items outside of their syntactic scope, and for their ability to interact with nondistributive event modifiers (Schein 1993, Champollion 2010a).

Lucas Champollion, New York City, June 25th, 2021

Contents

1

Linguistic applications of mereology

After an overview of the content to be covered, the course starts by a whirlwind tour of mereology, the singular-plural distinction, the count-mass distinction, higher-order properties, extensive and intensive measure functions, the telic-atelic opposition, and aspectual composition. This is a recap intended to refresh participants' memories; the corresponding topics are covered in more introductory materials such as the first half of the lecture notes at https://ling.auf.net/lingbuzz/002174 and the first few chapters of Champollion 2017.

1.1 Mereology

- **Mereology:** the study of parthood in philosophy and mathematical logic

- Used to formally represent the meaning of grammatical number and related phenomena

- In Figure 1.1, the small circles stand for Annie, Bonnie, Connie, our *atomic individuals*.

- The large circles are *sums*: formal objects that represent pluralities of children.

- The binary sum operation, ⊕, is taken to be associative, commutative, and idempotent.

- The lines indicate the parthood relation, ≤, a partial order (reflexive, transitive, and antisymmetric).

Figure 1.1: An algebraic structure
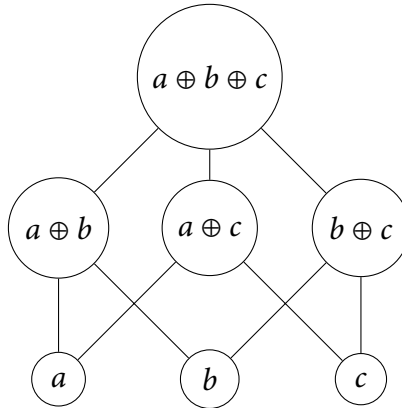
- Any nonempty set of things of the same sort (e.g. individuals, events, time intervals) has one and only one sum. This coincides with the operation of least upper bound.

- Models of classical extensional mereology (CEM) are essentially isomorphic to complete Boolean algebras with the bottom element removed, or equivalently complete semilattices with their bottom element removed (Tarski 1935, Pontow & Schubert 2006).

- CEM parthood is very similar to the subset relation (Table 1.1).

- Example: the powerset of a given set, with the empty set removed, and with the partial order given by the subset relation.

1.2 Selected literature

- – Textbooks:
    - ⋆ Montague-style formal semantics: Heim & Kratzer 1998
    - ⋆ Mathematical foundations: Partee, ter Meulen & Wall 1990

Table 1.1: Correspondences between CEM and set theory

	Property	CEM	Set theory
1	Reflexivity	$x \leq x$	$x \subseteq x$
2	Transitivity	$x \leq y \ \wedge \ y \leq z \to x \leq z$	$x \subseteq y \wedge y \subseteq z \to x \subseteq z$
3	Antisymmetry	$x \leq y \ \wedge \ y \leq x \to x = y$	$x \subseteq y \wedge y \subseteq x \to x = y$
4	Interdefinability	$x \leq y \leftrightarrow x \oplus y = y$	$x \subseteq y \leftrightarrow x \cup y = y$
5	Unique sum/union	$P \neq \varnothing \to \exists! z \ \mathrm{sum}(z, P)$	$\exists! z \ z = \bigcup P$
6	Associativity	$x \oplus (y \oplus z) = (x \oplus y) \oplus z$	$x \cup (y \cup z) = (x \cup y) \cup z$
7	Commutativity	$x \oplus y = y \oplus x$	$x \cup y = y \cup x$
8	Idempotence	$x \oplus x = x$	$x \cup x = x$
9	Unique separation	$x < y \to$ $\exists! z [x \oplus z = y \wedge \neg x \circ z]$	$x \subset y \to \exists! z [z = y - x]$

- ⋆ Algebraic semantics: Landman 1991

  - Books on algebraic semantics: Link 1998a, Landman 2000

  - Seminal articles on algebraic semantics: Landman 1996, Krifka 1998

  - Mereology surveys: Simons 1987, Casati & Varzi 1999, Varzi 2010

  - Linguistic applications of mereology: Champollion & Krifka 2016

  - This course is based on the book Champollion 2017

  - The book overlaps with but is not identical to Champollion 2010b

  - See champollion.com/book-parts-of-a-whole/ for excerpts and other information

1.3   Algebraic closure and the plural

- Algebraic closure closes any predicate (or set) *P* under sum formation:

(1)   **Definition: Algebraic closure (Link 1983)**
The algebraic closure $^*P$ of a set P is defined as $\{x \mid \exists P' \subseteq P[x = \oplus P']\}$.
(This is the set that contains any sum of things taken from P.)

(2)   a.   John is a boy.
b.   Bill is a boy.
c. $\Rightarrow$ John and Bill are boys.

(3) $\mathrm{boy}(j) \wedge \mathrm{boy}(b) \Rightarrow {}^*\mathrm{boy}(j \oplus b)$

(4)   **Definition: Algebraic closure for relations**
The algebraic closure $^*R$ of a non-functional relation R is defined as
$\{\vec{x} \mid \exists R' \subseteq R[\vec{x} = \oplus R']\}$
(The algebraic closure of a relation R is the relation that contains any sum of tuples each contained in R.)

(5)   **Definition: Algebraic closure for partial functions**
The algebraic closure $^*f$ of a partial function f is defined as
$\lambda x : x \in {}^*\mathrm{dom}(f). \oplus \{y \mid \exists z[z \leq x \wedge y = f(z)]\}$
(The algebraic closure of f is the partial function that maps any sum of things each contained in the domain of f to the sum of their values.)

- On both the inclusive and exclusive view, plural nouns are *cumulative*.

(6)   **Definition: Cumulative reference**
$\mathrm{CUM}(P) \stackrel{\mathrm{def}}{=} \forall x[P(x) \to \forall y[P(y) \to P(x \oplus y)]]$
(A predicate P is cumulative if and only if whenever it holds of two things, it also holds of their sum.)

- The property of cumulativity is common to plural nouns and mass nouns (see below).

Figure 1.2: Different views on the plural.
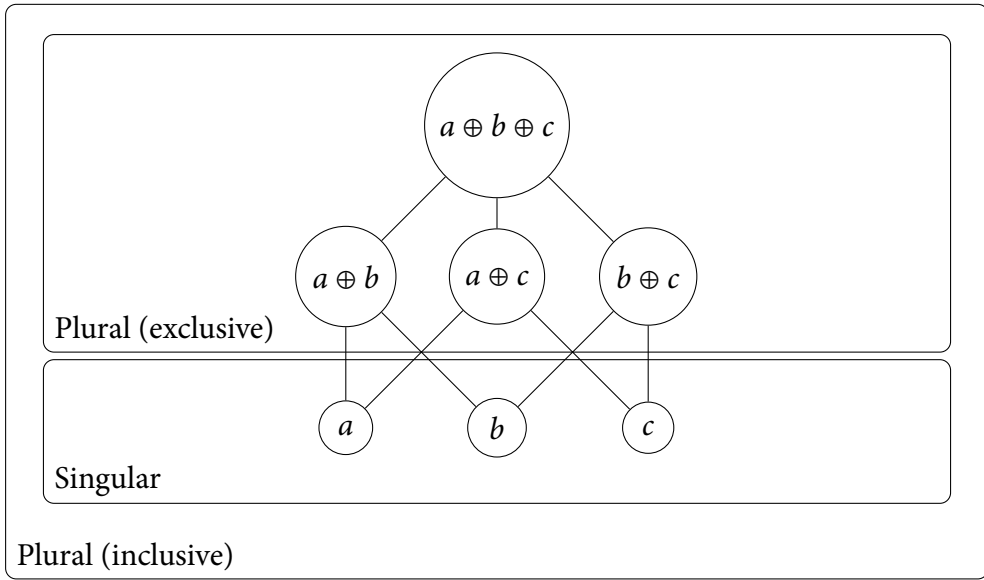
1.4   Singular count nouns

- Counting involves mapping to numbers. Let a "singular individual" be something which is mapped to the number 1, something to which we can refer by using a singular noun.

- One can assume that all singular individuals are *atoms*: the cat's leg is not a part of the cat.

    (7)   **Definition: Atom**
    $\text{Atom}(x) \stackrel{\text{def}}{=} \neg\exists y[y < x]$
    (An atom is something which has no proper parts.)

    (8)   **Definition: Atomic part**
    $x \leq_{Atom} y \stackrel{\text{def}}{=} x \leq y \ \wedge \ \text{Atom}(x)$
    (Being an atomic part means being atomic and being a part.)

(9)   **Definition: Quantized reference**
$\text{QUA}(P) \overset{\text{def}}{=} \forall x[P(x) \to \forall y[y < x \to \neg P(y)]]$
(A predicate P is quantized if and only if whenever it holds of something, it does not hold of any its proper parts.)

1.5   Mass nouns and atomicity

- Anything which can be referred to by a proper name, or denoted by using a common noun

- Objects form a mereology, so they include plural objects (sums)

  - Individuals: firemen, apples, chairs, opinions, committees

  - Substances: the water in my cup or the air which we breathe

- Mass nouns have cumulative reference: add water to water and you get water

- In this, they parallel plural nouns (Link 1983)

- Mass nouns were proposed to have divisive reference (Cheng 1973); but this position is no longer popular (minimal-parts problem)

(10)   **Definition: Divisive reference**
$\text{DIV}(P) \overset{\text{def}}{=} \forall x[P(x) \to \forall y[y < x \to P(y)]]$
(A predicate P is divisive if and only if whenever it holds of something, it also holds of each of its proper parts.)

1.6   Lexical cumulativity

- Many authors assume *lexical cumulativity*: whenever two events are in the denotation of a verb, so is their sum (Scha 1981, Schein 1986, 1993, Lasersohn 1989, Krifka 1989, 1992, Landman 1996, 2000, Kratzer 2008).

(11) a. John slept.
      b. Mary slept.
      c. ⇒ John and Mary slept.

(12) a. John saw Bill.
      b. Mary saw Sue.
      c. ⇒ John and Mary saw Bill and Sue.

- Verbs have plural denotations: they obey the same equation as plural count nouns on the inclusive view

(13) $[\![V]\!] = {}^*[\![V]\!]$

(14) $[\![N_{pl}]\!] = {}^*[\![N_{sg}]\!]$

- It is customary to indicate lexical cumulativity by writing $\lambda e[{}^*\mathrm{see}(e)]$ for the meaning of the verb *see* instead of $\lambda e[\mathrm{see}(e)]$.

  This entailment is parallel to the entailment from singular to plural nouns:

(15) a. John is a boy.
      b. Bill is a boy.
      c. ⇒ John and Bill are boys.

- Lexical cumulativity does not entail that all verb phrases have cumulative reference. For example, the sum of two events in the denotation of the verb phrase *carry exactly two pianos* is not again in its denotation, because it involves four rather than two pianos.

(16) **Cumulativity assumption for thematic roles**
      For any thematic role θ it holds that $\theta = {}^*\theta$. This entails that
      $\forall e, e', x, y[\theta(e) = x \;\wedge\; \theta(e') = y \;\rightarrow\; \theta(e \oplus e') = x \oplus y]$

- As a consequence of (16), thematic roles are homomorphisms with respect to the ⊕ operation:

(17) **Fact: Thematic roles are sum homomorphisms**
      For any thematic role θ, it holds that $\theta(e \oplus e') = \theta(e) \oplus \theta(e')$.
      (The θ of the sum of two events is the sum of their θs.)

Exercise 1.1 Does the verb phrase *see John* have cumulative reference? $\square$

1.7   Trace functions

- Various partial functions formalize the relationships between the domains
  of the ontological zoo, as shown in Figure 1.3

Figure 1.3: The world (some details omitted)
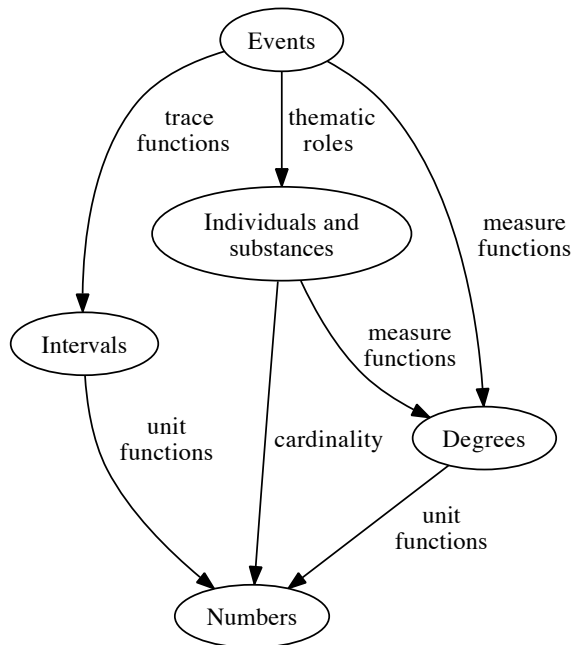
- Trace functions map events to intervals which represent their temporal and spatial locations

  - τ, the temporal trace or *runtime*
  - σ, the spatial trace

- Trace functions are sum homomorphisms (Krifka 1998), like thematic roles.

  (18)   **Trace functions are sum homomorphisms**
  σ is a sum homomorphism: $\sigma(e \oplus e') = \sigma(e) \oplus \sigma(e')$
  τ is a sum homomorphism: $\tau(e \oplus e') = \tau(e) \oplus \tau(e')$
  (The location/runtime of the sum of two events is the sum of their locations/runtimes.)

1.8   Measure functions and degrees

- While trace functions map entities to intervals, measure functions map entities to degrees (but some authors conflate them, e.g. Kratzer 2001)

- Typical measure functions: height, weight, speed, temperature

- Degrees are totally ordered quantities assigned by measure functions

1.9   Unit functions

- For Lønning 1987, degrees occupy an intermediate layer betwen individuals and numbers (see also Schwarzschild 2006).

- Measure nouns like *liter*, *kilogram*, *year* denote functions from degrees to numbers: what I will call *unit functions*.

  (19)   a. $[\![\text{liter}]\!] = [\![\text{liters}]\!] = \lambda n \lambda d [\text{liters}(d) = n]$

b.   ⟦year⟧ = ⟦years⟧ = $\lambda n \lambda t [\text{years}(t) = n]$

- Example: John weighs 150 pounds (68 kilograms) and measures six feet (183 centimeters). *Weight* and *height* are measure functions, *feet* and *centimeters* are unit functions

  (20)   a.   pounds(weight(john))=150
          b.   kilograms(weight(john))=68
          c.   feet(height(john))=6
          d.   centimeters(height(john))=183

- Advantage of Lønning's split: underspecifiation in pseudopartitives

  (21)   three inches of oil
          a. $\lambda x [\text{oil}(x) \wedge \text{inches}(\text{height}(x)) = 3]$ *(by height)*
          b. $\lambda x [\text{oil}(x) \wedge \text{inches}(\text{diameter}(x)) = 3]$ *(by diameter)*

- Ambiguity of container pseudopartitives (Rothstein 2009: and references therein):

  (22)   three glasses of wine
          a.   *Measure reading:* $\lambda x [\text{wine}(x) \wedge \text{glasses}(x) = 3]$
                (a quantity of wine that corresponds to three glassfuls)
          b.   *Individuating reading:* $\lambda x [|x| = 3 \wedge {}^{*}\text{glass}(x) \wedge \text{contains}(x, \text{wine})]$
                (three actual glasses containing wine)

1.10   Aspectual composition

- Predicates can be telic or atelic.

  – Atelic predicates: walk, sleep, talk, eat apples, run, run towards the store
     (≈ *as soon as you start X-ing, you have already X-ed*)

- – Telic predicates: build a house, finish talking, eat ten apples, run to the store
  (≈ *you need to reach a set terminal point in order to have X-ed*)

- Traditionally, atelicity is understood as the subinterval property or divisive reference. Telicity is understood as quantized reference. This brings out the parallel between the telic/atelic and count/mass oppositions (e.g. Bach 1986).

  (23)   a.   telic : atelic :: count : mass
         b.   quantized : (approximate) subinterval :: quantized : (approximate) divisive

- We will use the following definition of the subinterval property:

  (24)   SUBINTERVAL(P) $=_{def}$
         $\forall e[P(e) \rightarrow \forall i[i < \tau(e) \rightarrow \exists e'[P(e') \wedge e' < e \wedge i = \tau(e')]]]$
         (Whenever P holds of an event e, then at every subinterval of the runtime of e, there is a subevent of which P also holds.)

  (25)   *eat ten apples for three hours
         Failing presupposition: SUBINTERVAL($[\![$eat ten apples$]\!]$), i.e. every part of the runtime of an eating-ten-apples event e is the runtime of another eating-ten-apples event that is a part of e.

- *Aspectual composition* is the problem of how complex constituents acquire the telic/atelic distinction from their parts.(Verkuyl 1972, Krifka 1998)

- With "incremental theme" verbs like *eat*, the correspondence is clear:

  (26)   a.   eat apples / applesauce for an hour
         b.   *eat an apple / two apples / the apple for an hour

  (27)   a.   count : mass :: telic : atelic
         b.   apple : apples :: eat an apple : eat apples

  (28)   a.   drink wine for an hour

  b. *drink a glass of wine for an hour

- With "holistic theme" verbs like *push* and *see*, the pattern is different:

 (29) a. push carts for an hour
    b. push a cart for an hour

 (30) a. look at apples / applesauce for an hour
    b. look at an apple / two apples / the apple for an hour

- *Verkuyl's Generalization* (Verkuyl 1972): When the direct object of an incremental-theme verb is a count expression, we have a telic predicate, otherwise an atelic one.

- Krifka 1992: in incremental-theme verbs (also called "measuring-out" verbs, among other things), the parts of the event can be related to the parts of the theme (see Figure 1.4).

Figure 1.4: Incremental theme of *drink wine*, from Krifka 1992
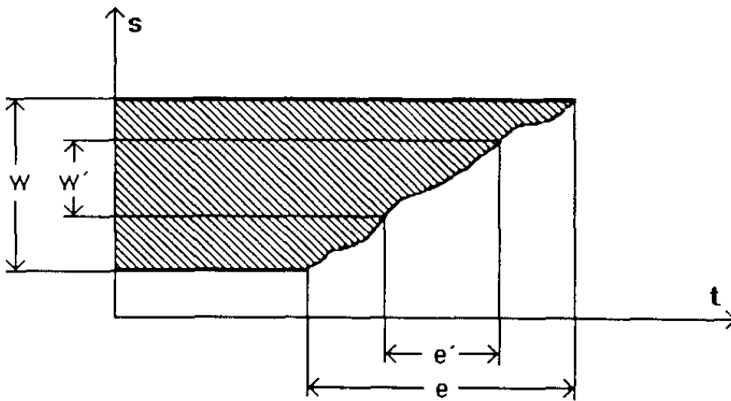
- Following Krifka, we can formalize the difference between holistic-theme and incremental-theme verbs by meaning postulates.

(31)   **Definition: Incrementality**
       Incremental$_\theta$(P) $\Leftrightarrow \forall e \forall e' \forall x [\theta(e) = x \wedge e' < e \rightarrow \theta(e') < x]$

(32)   **Definition: Holism**
       Holistic$_\theta$(P) $\Leftrightarrow \forall e \forall e' \forall x [\theta(e) = x \wedge e' < e \rightarrow \theta(e') = x]$

(33)   **Meaning postulates**
       a.   Incremental$_{theme}$($[\![$eat$]\!]$)
       b.   Incremental$_{theme}$($[\![$drink$]\!]$)
       c.   Holistic$_{theme}$($[\![$see$]\!]$)

- Then we apply these meaning postulates to prove or disprove that the various VPs above have divisive reference or the subinterval property.

- **Claim:** $[\![$eat two apples$]\!]$ does not have the subinterval property.

- **Proof:** Suppose it has, then let e be an event in its denotation whose runtime is an hour. From the definition of the subinterval property, (24), at each subinterval of this hour there must be a proper subevent of e whose theme is again two apples. Let e' be any of these proper subevents. Let the theme of e be x and the theme of e' be y. Then x and y are each a sum of two apples. From the "incremental theme" meaning postulate in (33a) we know that y is a proper part of x'. Since *two apples* is quantized, x and y can not both be two apples. Contradiction.

Exercise 1.2 Why does the proof not go through for *see two apples*? Why does it not go through for *eat apples*? □

2

Stratified reference: a bridge between distributivity, aspect, and measurement

This lecture introduces stratified reference, which provides a unified perspective on the domains of measurement, aspect, and distributivity. It explains the linguistic relevance of the difference between extensive measure functions like volume and intensive measure functions like temperature, as illustrated by the pseudopartitives thirty liters of water *vs.* *thirty degrees Celsius of water *(Krifka 1998, Schwarzschild 2006). Subsuming these previous accounts, stratified reference correctly predicts the monotonicity constraint: such constructions disallow measure functions that generally return the same value on an entity and on its parts. For example, in order for* *thirty degrees Celsius of water *to be acceptable, it would have to describe a water entity whose parts are colder than itself; but there are no such entities. Stratified reference relativizes unboundedness to just one dimension or measure function at a time. This makes it possible to account for examples like* five feet of snow *even though not every part of a five-foot snow layer of snow is less than five feet high.*

2.1   The measurement puzzle

Pseudopartitives reject some measure functions (Krifka 1998, Schwarzschild 2006)

(1)	a.	five pounds of rice	*weight*
	b.	five liters of water	*volume*
	c.	five hours of talks	*duration*

d.	five miles of railroad tracks	*spatial extent*
e.	*five miles per hour of driving	*speed*
f.	*five degrees Celsius of water	*temperature*

Several other constructions behave analogously:

(2)   more rope   *by length / by weight / *by temperature*

(3)   *five miles per hour of my driving   *speed*

2.1.1   Previous work

Schwarzschild 2006: Only *monotonic* measure functions are admissible.

- A measure function μ is *monotonic* iff for any two entities a and b, if a is a proper part of b, then $\mu(a) < \mu(b)$. (See also Krifka 1998.)

Examples:

- Volume is monotonic ⇝ *thirty liters of water*

- Temperature is not monotonic ⇝ **thirty degrees Celsius of water*

- What about height? It had better be monotonic: ⇝ *five feet of snow*

Problem: The snow that fell on West Berlin is a proper part of the snow that fell on Berlin. But, we don't conclude that the height of the snow in West Berlin was less than the snow that fell on Berlin. So height is **not monotonic**.

2.1.2   Novel observation

Measure functions rejected by pseudopartitives are also rejected by *for*-adverbials.

(4)	a.	John waited for five hours.	*duration*
	b.	The crack widens for five meters.	*spatial extent*
	c.	*John drove for thirty miles an hour.	*speed*
	d.	*The soup boiled for 100 degrees Celsius.	*temperature*

This connection allows us to tap into the literature on aspect.

Plan of this lecture:

- Introduce *stratified reference*, which generalizes the telic/atelic contrast.

- Derive the restriction on measure functions from this concept.

2.2   The aspect puzzle

As we've seen before, telicity is a property of predicates (Krifka 1998). But which one?

Classical answer   To be atelic means to have the *subinterval property* (e.g. Bennett & Partee 1972, Dowty 1979).

We have previously given this event-based version of the subinterval property:

(5)   SUBINTERVAL$(P) =_{def}$
$\forall e[P(e) \rightarrow \forall i[i < \tau(e) \rightarrow \exists e'[P(e') \wedge e' < e \wedge i = \tau(e')]]]$
(Whenever P holds of an event e, then at every subinterval of the runtime of e, there is a subevent of which P also holds.)

2.2.1   Problems with the subinterval property

First problem   The "minimal-parts problem" (Taylor 1977, Dowty 1979):

(6)   John and Mary waltzed for an hour
$\not\Rightarrow$ #John and Mary waltzed within every single moment of the hour
$\Rightarrow$ John and Mary waltzed within every short subinterval of the hour

The minimal length varies relative to the length of the bigger interval:

(7)   The Chinese people have created abundant folk arts … passed on from generation to generation for thousands of years.

Second problem   Spatial *for*-adverbials (Gawron 2005):

(8)   a.   The crack **widens** for 5 meters.   *spatially atelic*
        b.   #The crack **widens 2cm** for 5 meters.   *spatially telic*

(9)   a.   The road ends in a mile.
        b.   *The road ends for a mile.

(10)   a.   *The road meanders in a mile.
        b.   The road meanders for a mile.

(11)   The police blocked streets for miles around [the museum].[1]

Spatial and temporal *for*-adverbials impose different constraints – see Figure 2.1.

(12)   a.   John pushed carts to the store for fifty minutes.
        b.   John pushed carts to the store for fifty meters.
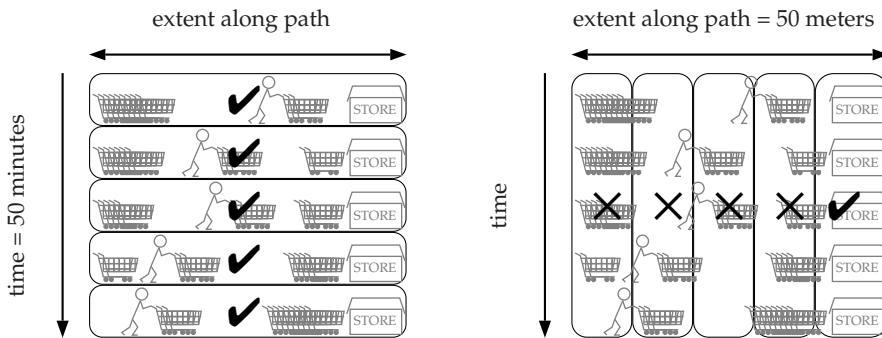
Figure 2.1: *John pushed carts to the store for fifty minutes* vs. *for fifty meters*

(13)   Snow fell throughout the area for two straight days.[2]

        a.   ⇒ Every part of "two straight days" is the runtime of an event in ⟦Snow fell throughout the area⟧
        b.   ⇏ Every part of "throughout the area" is the location of an event in ⟦Snow fell throughout the area⟧

---

[1] Attested example, New York Times, February 24, 2009. Thanks to Cleo Condoravdi.
[2] Attested example (http://community.lawyers.com/forums/t/17235.aspx).

(14)   Wine flowed from the jar to the floor for five minutes.[3]

    a.   ⇒ Every part of "five minutes" is the runtime of an event in ⟦Wine fell from the jar to the floor⟧

    b.   ⇏ Every part of "from the jar to the floor" is the location of an event in ⟦Wine fell from the jar to the floor⟧

2.2.2   Generalizing the subinterval property

> **What the subinterval property says:** An atelic predicate P distributes along the *time* dimension down to intervals of *infinitely short length*.
>
> **What it should say:** An atelic predicate P distributes along the _____ dimension down to intervals of _____ length.
>
> That is, we want to *parametrize* the subinterval property.

We start with applying the subinterval property to *waltz*:

(15) $\forall e[\text{waltz}(e) \rightarrow \forall i[i < \tau(e) \rightarrow \exists e'[\text{waltz}(e') \wedge e' < e \wedge i = \tau(e')]]]$
(Whenever *waltz* holds of an event e, then at every subinterval of the runtime of e, there is a subevent of which *waltz* also holds.)

We want to be able to say:

(16)   Whenever *waltz* holds of an event, there is a way of dividing this event into subevents with shorter runtimes such that *waltz* also holds of each of these subevents.

To express this formally, we use the star operator.

- Reminder: $x \in {}^*(\lambda y.B(y))$ means: x consists of one or more parts of which B holds

With the star operator, we can express the gist of (16) as follows:

---

[3] Beavers 2008

(17) $e \in {}^*\lambda e' \begin{pmatrix} \text{waltz}(e') \wedge \\ \tau(e') < \tau(e) \end{pmatrix}$

Let us say that *waltz* has **stratified reference** (SR) with respect to an event e, the dimension *runtime* and the granularity $\lambda t.\ t < \tau(e)$ just in case (17) above is true.

(18)   **Stratified reference (Example)**
Let   "$\text{SR}_{runtime,\ \lambda t.\ t < \tau(e)}(\lambda e[\text{waltz}(e)])(e)$"   abbreviate (17).

By abstracting from this example, we arrive at the following definition:

(19)   **Stratified reference (Definition)**
$\text{SR}_{f,g}(P)(x) \stackrel{\text{def}}{=} x \in {}^*\lambda y \begin{pmatrix} P(y) \wedge \\ g(f(y)) \end{pmatrix}$
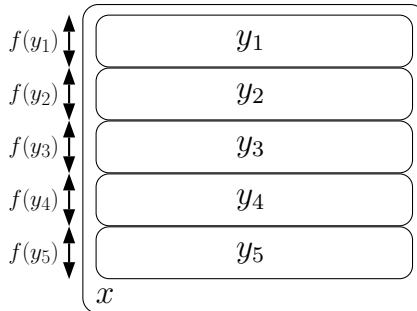
Figure 2.2: Illustration of stratified reference

- A predicate P stratifies x with respect to a function f and a granularity level g if and only if x can be exhaustively divided into parts which are each in P and which are each mapped by f to something in g.

The answer to the aspect puzzle.
Being atelic means having stratified reference with respect to time and a suitably instantiated granularity parameter.

For-adverbials presuppose stratified reference, not the subinterval property:

(20)   waltz for an hour
       Satisfied presupposition: $\text{SR}_{\tau,\ \lambda t.\ t < \tau(e)}(\llbracket\text{waltz}\rrbracket)(e)$

$$e \in {}^*\lambda e' \begin{pmatrix} e' \in \llbracket\text{waltz}\rrbracket \wedge \\ \tau(e') < \tau(e) \end{pmatrix}$$

The waltzing event e in question can be exhaustively divided into parts ("strata") which are also waltzing events and whose runtimes are properly included in the runtime of e.

(21)   eat apples for three hours
       Satisfied presupposition: $\text{SR}_{\tau,\ \lambda t.\ t < \tau(e)}(\llbracket\text{eat apples}\rrbracket)(e)$

$$e \in {}^*\lambda e' \begin{pmatrix} e' \in \llbracket\text{eat apples}\rrbracket \wedge \\ \tau(e') < \tau(e) \end{pmatrix}$$

The apple-eating event e in question can be exhaustively divided into apple-eating events whose runtimes are properly included in the runtime of e.

(22)   *eat ten apples for three hours
       Failing presupposition: $\text{SR}_{\tau,\ \lambda t.\ t < \tau(e)}(\llbracket\text{eat ten apples}\rrbracket)(e)$

$$e \in {}^*\lambda e' \begin{pmatrix} e' \in \llbracket\text{eat ten apples}\rrbracket \wedge \\ \tau(e') < \tau(e) \end{pmatrix}$$

The ten-apple-eating event e in question can be exhaustively divided into ten-apple-eating events whose runtimes are properly included in the runtime of e.

(23)   *widen 2cm for 5 meters
       Failing presupposition: $\text{SR}_{\tau,\ \lambda t.\ t < \sigma(e)}(\llbracket\text{widen 2cm}\rrbracket)(e)$

$$e \in {}^*\lambda e' \begin{pmatrix} e' \in \llbracket\text{widen 2cm}\rrbracket \wedge \\ \sigma(e') < \sigma(e) \end{pmatrix}$$

The 2-cm-widening event e in question can be exhaustively divided into 2-cm-widening events whose spatial locations are properly included in the spatial location of e.

2.3   Back to the measurement puzzle

Why can you not say *thirty degrees of water*?

As we have seen, *for*-adverbials reject certain measure functions too:

(24)   a.   *John drove for thirty miles an hour.   **speed*
       b.   *The soup boiled for 100 degrees Celsius.   **temperature*

Null assumption   These sentences have parametrized presuppositions of the same kind as temporal and spatial *for*-adverbials.

(25)   *drive for thirty miles per hour
       Failing presupposition: $\text{SR}_{speed,\ \lambda d.\ d\ <\ speed(e)}(\llbracket \text{drive} \rrbracket)(e)$

$$e \in {}^{*}\lambda e' \left(\begin{array}{l} e' \in \llbracket \text{drive} \rrbracket \ \wedge \\ \text{speed}(e') < \text{speed}(e) \end{array} \right)$$

The driving event e in question can be exhaustively divided into driving events whose speeds are less than the speed of e.

(26)   *boil for 100 degrees Celsius
       Failing presupposition: $\text{SR}_{temperature,\ \lambda d.\ d\ <\ temperature(e)}(\llbracket \text{boil} \rrbracket)(e)$

$$e \in {}^{*}\lambda e' \left(\begin{array}{l} e' \in \llbracket \text{boil} \rrbracket \ \wedge \\ \text{temperature}(e') < \text{temperature}(e) \end{array} \right)$$

The boiling event e in question can be exhaustively divided into boiling events whose temperatures are less than the temperatures of e.

Now we transfer this idea to pseudopartitives.
      Intuition: *run for three hours ≈ three hours of running*

(27)   a.   five pounds of **books**   *plural*
       b.   thirty liters of **water**   *mass*
       c.   *five pounds of **book**   **singular*

"John walked for three hours." | three hours | …is the | runtime | of … | walk |
"three hours of walking" | three hours | …is the | runtime | of … | walk |
"three liters of water" | three liters | …is the | volume | of … | water |

2.3.1   Baseline examples

Assumption: Same presuppositions for *for*-adverbials and pseudopartitives.

(28)   run for three hours / three hours of running
       Satisfied presupposition: $\text{SR}_{\tau,\,\gamma(\tau,\,e)}(\llbracket\text{run}\rrbracket)(e)$
       The running event e in question can be exhaustively divided into running events whose runtimes are properly included in the runtime of e.

The dimension parameter is the appropriate measure function.

(29)   thirty liters of water
       Satisfied presupposition: $\text{SR}_{volume,\,\gamma(volume,\,x)}(\llbracket\text{water}\rrbracket)(x)$
       The portion of water x in question can be exhaustively divided into portions of water whose volumes are lower than the volume of x.

2.3.2   Temperature in pseudopartitives

No smaller temperatures as you go from bigger to smaller amounts of substance.

(30)   *thirty degrees Celsius of water
       Failing presupposition: $\text{SR}_{temperature,\,\gamma(temperature,\,x)}(\llbracket\text{water}\rrbracket)(x)$
       The portion of water x in question can be exhaustively divided into portions of water whose temperatures are lower than the temperature of x.

2.3.3   The problematic snow example

Unlike Schwarzschild's, this account has no monotonicity requirement.

(31)   five feet of snow
       Satisfied presupposition: $\text{SR}_{height,\,\gamma(height,\,x)}(\llbracket\text{snow}\rrbracket)(x)$
       The portion of snow x in question can be exhaustively divided into portions of snow whose heights are lower than the height of x.

2.3.4   Ruling out singular count nouns

Singular count nouns are ruled out because they are quantized.

(32)   *five pounds of book
       Failing presupposition: $\text{SR}_{weight,\,\gamma(weight,\,x)}(\llbracket\text{book}\rrbracket)(x)$

Figure 2.3: Accepting *five feet of snow*
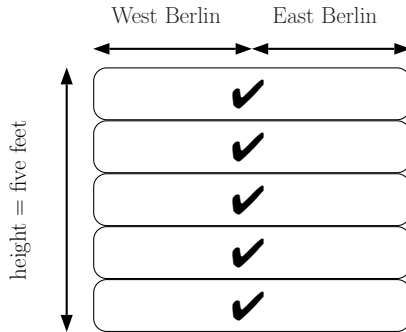
The book x in question can be exhaustively divided into books whose weights are lower than the weight of x.

The answers to the measurement puzzle.

1. How can we characterize the class of admissible measure functions?

   - A pseudopartitive has to satisfy stratified reference, where the dimension parameter is specified by the measure function.

2. Why are not all measure functions admissible in the first place?

   - The constraint on measure functions is also instantiated in *for*-adverbials and other constructions.

2.4   What is distributivity?

- Most commonly: Property of predicates, as opposed to collectivity

  - *Distributive:* Walk, smile, take a breath
  - *Collective:* Meet, gather, be numerous

- *Each* diagnoses the distributive/collective opposition (e.g. Dowty & Brodie 1984, Link 1987)

(33) a. Three boys each **read a book**.   *distributive*
      b. *Three boys each **met**.   **collective*

- Also often understood as a property of quantifiers

  – *Distributive:* Each man, every woman, all the men

  – *Nondistributive:* some man, a man, three women

2.5   Distributivity and stratified reference

2.6   Introduction

- Stratified theory allows us to capture cross-categorial parallels between atelic aspect, mass reference, plural reference, and distributivity.

- There is an intuitive parallelism between the telic-atelic, collective-distributive, singular-plural, and count-mass oppositions.

- Singular, telic, and collective predicates are delimited or bounded in ways that plural, mass, atelic, and distributive predicates are not.

- The **boundedness question:** How can the difference between boundedness and unboundedness be formally characterized?

- The use of the word *distributivity* generally indicates the application of a predicate to the members or subsets of a set, or to the parts of an entity.

- There are no standard definitions of distributivity. But it is diagnosed by the presence of *distributive entailments*.

(34) **Distributive predicates**
      a. The children smiled. ≈ Every child smiled.
      b. John and Mary sang. ≈ John sang and Mary sang.

(35)   **Collective predicates**
a.   The children gathered. ⊭ *Every child gathered.
b.   John and Mary met. ⊭ *John met and Mary met.

- Two constituents that contribute to the content of a distributive entailment stand in a distributive relation.

- A theory that relies heavily on this concept is developed in Choe 1987.

(36)   a.   Al and Bill each ate a pizza.
b.   Al and Bill ate a pizza.

- In (36a), the relation between the subject and the verb phrase is obligatorily distributive. In (36b), it is optionally distributive.

- A distributive relation can be indicated by distributive markers such as *each* in (36a).

2.7   Distributive constructions

- We can also see distributivity as a property of entire constructions.

- A *distributive construction* is a lexicosyntactic configuration that imposes an obligatory distributive relation between two of its constituents.

- Sentences with *each* are distributive constructions.

2.7.1   Pseudopartitives

- The pseudopartitive can also be classified as a distributive construction.

(37)   Three liters of water are sufficient.

- Distributive entailment: among the parts of the water in question there exist one liter of water, two liters of water, and so on.

- It is required by the construction, because every pseudopartitive gives rise to similar entailments.

- Since pseudopartitives obligatorily involve a distributive relation, I classify them as distributive constructions.

2.7.2 *For*-adverbials

- *For*-adverbials license entailments of the following kind:

(38) a.   John ran for five minutes.
      b.   ⇒ John ran for four minutes.
      c.   ⇒ John ran for three minutes.

- These entailments are distributive entailments because they are jointly determined by the *for*-adverbial and the predicate it modifies.

- So, *for*-adverbials are distributive constructions.

2.8 The components of a distributive relation

(39)   The boys (each) took a breath.

- There are many names for the components of a distributive relation. Table 2.1 lists a few.

- I adopt the terms *Key* and *Share*.

- The term *Share* refers to the constituent whose denotation is distributed over the parts of the referent of the other constituent.

- This other constituent is called the *Key*.

Table 2.1: Terms for the components of a distributive relation

Author	Name for *"The boys"*	Name for *"(took) a breath"*
Link 1998b	distributional domain	Distributive Share (DstrShr)
Choe 1987	Sorting Key (SrtKy)	Distributed Share (DstrShr)
Safir & Stowell 1988	Range NP	Distributing NP (DistNP)
Gil 1989, Choe 1991	Key	Share
Zimmermann 2002	DistKey	DistShare
Blaheta 2003	Dist phrase	Range
This work	**Key**	**Share**

- The property of reading a book is distributed over the individual boys, the property of being water is distributed over the liters, and the property of being an event of pushing a cart is distributed over the hours (see Table 2.2).

Table 2.2: A bridge from distributivity to aspect and measurement

Construction	Example	Key	Share	Map
Adverbial *each*	Three boys each laughed	three boys	laugh	agent
For-adverbial	John ran for three hours	three hours	John run	runtime
Pseudopartitive	three liters of water	three liters	water	volume

- According to our background assumptions, constituents are related by certain covert functions: thematic roles, trace functions, and measure functions.

- These functions coincide with distributive relations in a particular way: they always map entities associated with the Share to entities associated with the Key. I call tehm *Maps*.

2.9 The constraints on distributive constructions

- Distributive constructions impose constraints on their Shares:

(40) a. The boys each **walked**. *distributive*
b. *The boys each **met**. **collective*

(41) a. John **ran** for five minutes. *atelic*
b. *John **ran to the store** for five minutes. **telic*

(42) a. thirty pounds of **books** *plural*
b. thirty liters of **water** *mass*
c. *thirty pounds of **book** **singular*

- The following subsumes the traditional accounts of these constraints:

(43) **Distributivity Constraint**
A distributive construction whose Share is S, whose Map is m, and which is used to describe an entity x is acceptable only if S stratifies x with respect to m and a granularity level G specified by the construction (formally: $\mathrm{SR}_{m,G}(S)(x)$).

ex. **Definition: Stratified reference**
Let m (a "map" or "dimension") be any function from entities of type α to entities of type β, and let G (a "granularity level") be any predicate of entities of type β. Let S range over predicates of type $\langle \alpha, t \rangle$ where α is either e or v, and let x range over entities of type α. Then:

$$\mathrm{SR}_{m,g}(S)(x) \stackrel{\mathrm{def}}{=} x \in {}^{*}\lambda y \left(\begin{array}{c} S(y) \wedge \\ G(m(y)) \end{array} \right)$$

(A predicate S (the "Share") stratifies x with respect to a function m (the "dimension" or "map") and a predicate G (the "granularity level") if and only if x can be exhaustively divided into parts ("strata") which are each in m and which are each mapped by m to something in G.)

- The **boundedness question:** How can the difference between boundedness and unboundedness be formally characterized?

- My answer to the boundedness question: Stratified reference characterizes what it means to be an unbounded predicate. Unboundedness can be understood in more than one way. These different ways correspond to the parameters of stratified reference.

  - One and the same verb phrase, *push carts to the store*, can be distributive (unbounded with respect to agents) and telic (bounded with respect to runtime).

  - One and the same verb, *kill*, can be collective (bounded) with respect to agents and distributive (unbounded) with respect to themes.

    (44) a.   The police officers killed the two outlaws.
          b.   ⇒ The first outlaw was killed.
          c.   ⇏ The first police officer killed someone.

  - One and the same noun, *cable*, can be divisive along its length (bounded with respect to length) and nondivisive along its diameter (unbounded with respect to diameter).

3

Collectivity, cumulativity and *all*

This lecture develops an account for differences within the class of collective predicates, as exemplified by the contrast between all the students gathered *and* *all the students were numerous *(Dowty 1987, Winter 2001), and for the limited ability of* all *to take part in cumulative readings. Stratified reference is used to formulate meaning postulates that capture the fact that predicates like* gather *give rise to distributive inferences to subgroups, and to formulate the semantics of* all *in terms of a subgroup distributivity requirement.*

3.1 Cumulativity and *all*

- This is the stock example of a cumulative (scopeless) reading:

  (1) 600 Dutch firms use 5000 American computers. (Scha 1981)
       Cumulative reading: 600 Dutch firms each use at least one American computer and 5000 American computers are each used by at least one Dutch firm.

- Cumulative readings also occur with definite plurals:

  (2) The men in the room are married to the girls across the hall. (Kroch 1974)
       Cumulative reading: Each man in the room is married to a girl across the hall, and each girl across the hall is married to a man in the room.

- Zweig 2008, 2009 notes that *all* cannot give rise to cumulative readings:

(3) a. Three safari participants saw thirty zebras.
         Available cumulative reading: Three safari participants saw at least one zebra each, and thirty zebras were seen overall.
     b. All the safari participants saw thirty zebras.
         Unavailable cumulative reading: Each safari participant saw at least one zebra, and thirty zebras were seen overall.
         Available distributive reading: Each safari participant saw thirty zebras.

- Exception: dependent-plural readings, which can be seen as cumulative (Zweig 2008):

(4) a. Three safari participants saw zebras.
         Available cumulative reading: Three safari participants saw at least one zebra each, and at least two zebras were seen overall.
     b. All the safari participants saw zebras.
         ***Available** cumulative reading:* Each safari participant saw at least one zebra, and at least two zebras were seen overall.

- The word *each* is never compatible with cumulative readings but only with distributive readings:

(5) a. Each safari participant saw thirty zebras.
         Unavailable cumulative reading: Each safari participant saw at least one zebra, and thirty zebras were seen overall.
     b. Each safari participant saw zebras.
         Unavailable cumulative reading: Each safari participant saw at least one zebra, and at least two zebras were seen overall.

- **Dependent Plural Puzzle:** Why can *all* license dependent plurals but not *each*? (Answer in Zweig 2008, extended in Champollion 2010b. Not today.)

> **Thirty-Zebras Puzzle:** What is the relevant semantic distinction between *see zebras* and *see thirty zebras* so that the former is more permissive wrt. cumulative readings?

3.2   Collectivity and *all*

- *All* is incompatible with some collective predicates (Kroch 1974, Dowty 1987), which I will call *numerous-type*:

  (6)   a.   The students who came to the rally are numerous.
         b.   The boys surrounded the table.
         c.   The soldiers in this bataillon sufficed to defeat the army.

  (7)   a.   *All the students who came to the rally are numerous.
         b.   *All the boys surrounded the table.
         c.   *All the soldiers in this bataillon sufficed to defeat the army.

- But it is compatible with others, which I will call *gather-type*:

  (8)   a.   The students gathered in the hallway.
         b.   The professors met in the garden.
         c.   The soldiers dispersed.

  (9)   a.   All the students gathered in the hallway.
         b.   All the professors met in the garden.
         c.   All the soldiers dispersed.

  The word *each* is incompatible with any of these.

  (10)   a.   *Each student who came to the rally is numerous.
         b.   *Each boy surrounded the table.
         c.   *Each student gathered in the hallway.
         d.   *Each professor met in the garden.   etc.

Numerous-Gather Puzzle: What is the relevant semantic distinction between *be numerous* and *gather* so that only the latter is compatible with *all*?

Each-All Puzzle: What is the relevant semantic distinction between *each* and *all* so that only the latter is compatible with *gather*-type predicates?

- We will look for an answer in *for*-adverbials and the telic/atelic opposition.

(11)   a.   John **talked** for ten minutes.   *atelic*
        b.   *John **finished talking** for ten minutes.   *telic*

(12)   a.   John **ate apples** for ten minutes.   *atelic*
        b.   *John **ate three apples** for ten minutes.   *telic*

(13)   a.   John **wore yellow neckties at night** for a week.   *dependent plural ok*
        b.   #John **wore two yellow neckties at night** for a week.   *funny*

> **Aspect Puzzle**: What is the relevant semantic distinction between *eat apples* (atelic) and *eat three apples* (telic) so that only the former is compatible with *for*-adverbials?

- A *for*-adverbial cannot enter cumulative relation with an indefinite:

(14)   a.   John saw thirty zebras for three hours.
              Unavailable cumulative reading: John saw a total of thirty zebras over the course of a three-hour timespan.
        b.   John saw thirty zebras in three hours.
              Available cumulative reading: John saw a total of thirty zebras over the course of a three-hour timespan.

3.3   Explaining the similarities between *for* and *all*

- As we have seen in Lecture 2, *for*-adverbials work like this:

(15)   **Presupposition of *for three hours*:** The VPing event consists of one or more VPing events whose runtimes are shorter than its own.

– *John talked for three hours* requires something **true**: that the talking event consist of talking events whose runtimes are shorter than its own.

– **John finished talking for three hours* requires something **false**: that the finish-talking event consist of finish-talking events whose runtimes are shorter than its own.

- Claim: *all* imposes a constraint on the verb phrase predicate which is analogous to the presupposition of *for*-adverbials, except that it is universally quantified and the "dimension" involved is not runtime but the thematic role of the *all*-phrase, usually *agent*. (Cf. Champollion 2010b, Kuhn 2014, 2020, Dobrovie-Sorin 2014)

- Assume that *all* distributes the VP down to sums that are small in number.

(16)   **Presupposition of *all*:** Every VPing event consists of one or more VPing events whose agents are small in number.

– *All the children smiled* presupposes that every smiling event consists of one or more smiling events whose agents are small in number.

- What counts as small in number is determined by an irreflexive placeholder predicate ε that takes a comparison class K (either a set or a sum of entities) and an entity x, and returns true just in case x is small (typically small in number) compared to K.

- For example, $\varepsilon(\bigoplus \text{boy})(x)$ holds just in case x is small in number compared to the sum of all boys.

- I assume that ε is downward monotonic on its second argument (if x is small compared to K, then any part of x is also small compared to K).

(17) $\llbracket \text{all} \rrbracket = \lambda\theta\lambda y\lambda V\lambda e : \text{SR}_{\theta,\varepsilon(\theta(e))}(V).[V(e) \wedge {}^{*}\theta(e) = y]$

- *All* requires distributivity (Dowty 1987):

(18) a. All the boys smiled. ⇒ Each boy smiled.
      b. All the juries returned a guilty verdict. ⇒ Each jury did.
      c. The juries returned a guilty verdict. ⇏ Each jury did.

For *All the boys smiled*, this entry yields the following meaning:

(19) $\exists e : SR_{*agent,\varepsilon(*agent(e))}(*smile).[*smile(e) \wedge *agent(e) = \oplus boy]$
      (There is a smiling event e whose agents sum up to the boys, with
      the presupposition that *smile* has stratified reference along the
      agent dimension with granularity $\varepsilon(*agent(e))$.)]

Baseline example:

(20) All the children smiled.
      Presupposition: $SR_{*agent,\varepsilon(*agent(e))}(*smile)$
      $= \forall e'. *smile(e') \rightarrow e' \in *\lambda e'' \left(\begin{array}{c} *smile(e'') \wedge \\ \varepsilon(*agent(e))(*agent(e'')) \end{array} \right)$
      (*Smile* has stratified reference along the *agent* dimension with gran-
      ularity $\varepsilon(*agent(e))$ ("small in number compared to the agent of
      e") if and only if any smiling event can be divided into one or more
      smiling events whose agents are each small in number compared
      to the agent of e. This will be true because *smile* distributes.)

(21) a. All the safari participants saw thirty zebras.   *cumulative*
      b. All the safari participants saw zebras.   ✓ *cumulative*

We can rule out the cumulative reading of (21a) as a presupposition failure:

(22) **Failing presupposition:** $SR_{*agent,\varepsilon(*agent(e))}(\llbracket \text{see thirty zebras} \rrbracket)$
      (Every see-thirty-zebras event consists of subevents with small
      numbers of people as agents and in each of which thirty zebras are
      seen.)

The cumulative reading of (21b) is available, though:

(23)   **Satisfied presupposition:** $\text{SR}_{*\text{agent}, \varepsilon(*\text{agent}(e))}(\llbracket\text{see zebras}\rrbracket)$
(Every event in which at least one zebra is seen consists of subevents with small numbers of people as agents and in each of which at least one zebra is seen.)

Aspect Puzzle: What is the relevant semantic distinction between *eat apples* (atelic) and *eat three apples* (telic) so that only the former is compatible with *for*-adverbials?

  – Only *eat apples* has stratified reference with respect to time

Thirty-Zebras Puzzle: What is the relevant semantic distinction between *see zebras* and *see thirty zebras* so that the former is more permissive wrt. cumulative readings?

  – Only *see zebras* has stratified reference with respect to agents.

3.4   Explaining the behavior of *numerous* and *gather*

- Kuhn 2014, 2020: *gather* has stratified reference; *numerous* doesn't

(24)   **The subgroup distributivity hypothesis**
Gather-type predicates are those that distribute down to subgroups of small cardinality. Informally, if a collective predicate holds of a plural entity X, it will be a *gather*-type predicate just in case there is a way to divide X into at least two small, possibly overlapping subgroups (usually pairs or triples) such that the predicate applies to each of these subgroups.

3.4.1   *All* distinguishes between *be numerous* and *gather*

(25)   a.   *All the boys were numerous / surrounded the table.

b.   All the boys gathered / met / held hands / dispersed.

- Presupposition of (25a):

(26)   Every event *e* in the denotation of *be numerous* can be divided into one or more parts each of which is in the denotation of *be numerous* and has a small number of people as an agent.

- This fails because these parts don't qualify as *be numerous*.

- What about sentences with group nouns?

(27)   The enemy armies were numerous.   ✓ *distributive,* ✓ *collective*

(28)   All the enemy armies were numerous.   ✓ *distributive, *collective*

- Questions:

    – Why can *all* and *be numerous* cooccur?
    – Why does only (27) but not (28) have a distributive reading?

- Answer: because a collective predicate like *be numerous* can be shifted into a distributive reading

- In (27), the distributive shift (which I write D) leads to a distributive reading and its absence to a collective reading.

- ⟦be numerous⟧ = true of any event *e* whose agent is numerous (large in number)

- ⟦D(be numerous)⟧ = true of any event *e* which consists of one or more events that are in *be numerous* and whose agent is a pure atom (a singular individual).

- This predicate applies to fewer events than *be numerous* does.

- Example:

(29)   a.   The boys are numerous.   *true only if D is absent*
       b.   The army is numerous.   *true whether D is present or not*

- – We assume that *the boys* refers to a sum but *the army* refers to a pure atom (Barker 1992, Schwarzschild 1996).

- $D(\llbracket\text{be numerous}\rrbracket)$ satisfies the presupposition of *all* in (26), even though *be numerous* does not.

- Answers to the previous questions:

  - – Sentences like (27) are acceptable because they contain the distributive shift.

  - – They only have a distributive interpretation because this shift introduces distributivity.

- Kuhn's observation (see also Winter 2001): Gather-type predicates have stratified reference down to small numbers of people. Whenever a plurality of people gathers, any subgroup of them also gathers.

(30)   All the boys gathered.
       Presupposition: Every event e in the denotation of *gather* can be divided into one or more parts each of which is in the denotation of *gather* and has a small number of people as an agent.

3.4.2   *Gather* distinguishes between *each* and *all*

- Why are *gather*-type collective predicates incompatible with *all* but compatible with *each*?

(31)   a.   All the students gathered.
       b.   *Each student gathered.

- Idea:

- – *Every* and *each* distribute over events whose agents are atoms (individual people)

- – *All* distributes over events whose agents must be small in number but need not be atomic

- The presupposition of *each* is as follows:

  (32)   Every event in VP can be divided into parts which are in VP and whose agents are atoms.

- The presupposition of *all* is as follows:

  (33)   Every event in VP can be divided into parts which are in VP and whose agents are small numbers of people.

- *Gather* satisfies the presupposition of *all* but not of *each*.

> **Numerous-Gather Puzzle:** What is the relevant semantic distinction between *be numerous* and *gather* so that only the latter is compatible with *all*?
>
> - – Only gather-type collective predicates have stratified reference.
>
> **Each-All Puzzle:** What is the relevant semantic distinction between *each* and *all* so that only the latter is compatible with *gather*-type predicates?
>
> - – *Each* requires stratified reference down to atoms; *all* requires stratified reference down to entities that are small but not necessarily of cardinality one.

3.5   Conclusion

- *All* is an "almost distributive" determiner: it distributes but not all the way down.

- *Gather*-type predicates are "a bit distributive"; *numerous*-type predicates aren't.

3.6   Some open problems

- Some contexts improve the ability of certain *numerous*-type predicates to appear with *all*.

(34)   a.   Some of the boys were crying, but eventually (and after much discussion), all the boys formed a (nice) pyramid.
       b.   There was a lot of discussion, but eventually, all the boys decided unanimously to skip class.
       c.   I know it sounds kind of crazy but in fact all the weapons in this little village would suffice to defeat the US Army.
       d.   Some differences in acceptability and interpretation arise in connection with predicates like *build a raft* and *perform Hamlet*.

(35)   a.   It was a great evening. Some of the teachers played some early 20th century music, the others staged The Turn of the Screw and all the students performed Hamlet.

collective possible

       b.   A: So how was your class today?
            B: Great! All the students (in my class) performed Hamlet.

only (?) distributive

- These judgments appear to be variable. Some speakers I have consulted judge the collective reading to be available in (35b).

- These predicates appear to tolerate collective readings with *all* for some speakers even if they do not license subgroup distributivity. This category also includes predicates modified by adverbial *together* or by *same*:

(36)   a.   All the boys built a raft together. (Brisson 2003)
       b.   All the boys built a raft in the same room. (Jeremy Kuhn, p.c.)

- Adnominal *together* can prevent *all* from shifting collective predicates that lack subgroup distributivity into distributive predicates. The following

examples all lose their collective interpretations and acquire distributive interpretations when the word *together* is removed:

(37)   Jacob's sons each became the head of a tribe, and all the tribes together were called Israel.[4]

(38)   We gotta stick together: credit unions, leagues, and CUNA; all the elements together are unbeatable.[5]

(39)   All the data together are pretty compelling.[6]

- Possible explanation: a silent event predicate 'DO' that neutralizes the effect of *all* (Brisson 2003); adnominal *together* inserts some predicate similar to DO.

(40)   [All the boys [D_{agent} DO]] [built a raft].   *collective*

- But this overgenerates cumulative readings in sentences involving *all* and activity predicates:

(41)   All the linguistics majors { together / between them } dated five chemistry majors.

The cumulative reading disappears when adnominal *together* or *between them* is removed (Zweig 2009).

- Another problem is raised by the ability of *all* to license cumulative readings in passive sentences:

(42)   All the games were won by the Fijians and the Peruvians.

---

[4] Attested example, accessed May 25, 2016. http://twelvetribes.com/articles/why-twelve-tribes.
[5] Attested example, accessed May 25, 2016. https://www.ohiocreditunions.org/Advocacy/Pages/default.aspx.
[6] Attested example, accessed May 25, 2016. http://www.du.edu/give/faculty-linseman.html.

- Here, a possible explanation is that the by-phrase is adjoined at sentence level and remains outside of the scope of *all the games*, and therefore outside of the stratified-reference presupposition.

- Turning to *for*-adverbials, a cumulative reading seems to be available in examples like the following:

  (43)   a.   The search committee interviewed six job candidates for three hours.
          b.   Mary baked eight pans of cookies for four hours.[7]

- In a pilot study carried out via TurkTools (Erlewine & Kotek 2016), native speakers rated (43a) significantly lower as an appropriate description of a scenario where the interviews take place in sequence than as a description of a scenario when they all take place simultaneously. However, the absolute ratings were high, and the same contrast could not be reproduced for (43b).[8]

- Since these types of examples violate stratified reference on their cumulative reading, there is no way to derive it on the present account or on any algebraic account I know.

---

[7]Adapted from an attested example, accessed May 25, 2016. http://lulwanda2011.blogspot.com/2011/07/teacher-conference-ending.html.

[8]I am grateful to Hanna Muller and Linmin Zhang for their help with this pilot study.

4

Covert distributivity

Building on Champollion 2016a, this lecture considers how verb phrases such as build a raft *optionally acquire a distributive interpretation, and reformulates the covert distributivity operators of Link 1983 (the atomic D operator) and Schwarzschild 2006 (the nonatomic Part operator) in terms of the two parameters of stratified reference. By varying the granularity parameter, the difference between atomic and nonatomic views of distributivity is captured and clarified. By varying the dimension parameter, these distributivity operators are extended to the temporal domain and used to explain why indefinites in the syntactic scope of* for-*adverbials tend not to covary with them (?*John found a flea on his dog for a month, *Zucchi & White 2001).*

4.1   Introduction

- **What is distributivity?** In this lecture: a property of predicates

    - *Distributive:* e.g. walk, smile, take a breath (applies to a plurality just in case it applies to each of its members)
    - *Collective:* e.g. be numerous, gather, suffice to defeat the army (may apply to a plurality even if it does not apply to each of its members)

- Literature: Roberts 1987; Winter 2001, Section 6.2; Schwarzschild 1996, Chapter 6; Link 1997, Section 7.4; review articles: Champollion 2019, 2021.

4.2   Lexical and phrasal distributivity

(1)   **Lexical distributivity/collectivity** involves lexical predicates

a. The children smiled.   *distributive*
b. The children were numerous.   *collective*

(2) **Phrasal distributivity/collectivity** involves complex predicates

a. The girls are wearing a dress.   *distributive*
b. The girls are sharing a pizza.   *collective*
c. The girls are building a raft.   *collective/distributive*

- The difference between lexical and phrasal distributivity corresponds to the difference between what can and what cannot be described using meaning postulates

(3) **Meaning postulate: *smile* is distributive**
$\forall e[\text{smile}(e) \rightarrow e \in {}^*\lambda e'(\text{smile}(e') \wedge \text{Atom}(\text{agent}(e')))]$
(Every smiling event consists of one or more smiling events whose agents are atomic.)

- Meaning postulates can only apply to words. We cannot formulate a meaning postulate that says that *wear a dress* is distributive.

- Problems:

  - Meaning postulates are taken to be available only for lexical items
  - For mixed predicates like *build a raft*, we would need optional meaning postulates

- The classical solution is due to Link 1983: A covert distributive operator D adjusts the meaning of a verb phrase like *wear a dress* into *be a sum of people who each wear a dress*.

- D is in the lexicon, so it can apply to entire VPs (Dowty 1987, Roberts 1987, Lasersohn 1995).

- Link's D operator introduces a universal quantifier:

(4) $[\![\text{D}^{Link}]\!] = \lambda P_{\langle e,t\rangle}\lambda x\forall y[y \leq_{Atom} x \rightarrow P(y)]$
(Takes a predicate P over individuals and returns a predicate that applies to any individual whose atomic parts each satisfy P.)

(5)   a.   The girls built a raft.
              ≈ The girls built a raft together.   *collective*
        b.   The girls D^{Link}(built a raft).
              ≈ The girls each built a raft.   *distributive*

- This allows us to model the distributive meaning of (2a):

(6)   The girls D(are wearing a dress.)
        $\forall y[y \leq_{Atom} \bigoplus girl \rightarrow \exists z[\text{dress}(z) \wedge \text{wear}(y, z)]]$
        (Every atomic part of the sum of all girls wears a dress.)

- Based on earlier work by Eddy Ruys, Winter 2001 observes that the existential and the distributivity imports of numeral indefinites can have two distinct scopes.

(7)   If three workers in our staff have a baby soon we will have to face
        some hard organizational problems.
        a.   If any three workers have a baby, there will be problems. if >
               $3 > D > 1$
        b.   There are three workers such that if each of them has a baby,
               there will be problems. $3 > \text{if} > D > 1$

- Unlike the indefinite, the distributive operator cannot take scope outside
  of the *if*-island:

(8)   a.   *There are three workers such that for each x of them, if x has
               a baby, there will be problems. $3 > D > \text{if} > 1$

4.2.1   Reformulating the D operator

- Link's formulation of the D operator needs to be adjusted for several
  reasons:

  – If we assume with Landman 1996 that groups are atoms too ("impure"
    atoms) and that *the girls* can introduce a group, then we need to

specify that D distributes over "pure" atoms (singular individuals) only.

– If VPs are of type $\langle v, t \rangle$ instead of $\langle e, t \rangle$, we need to repair the type mismatch.

– We also need to be able to coindex D with different thematic roles (Lasersohn 1995).

(9)   a.   The first-year students D(took an exam). *Target: agent*
       b.   John D(gave a pumpkin pie) to two girls.   *Target: recipient*
       c.   John D(summarized) the articles.   *Target: theme*

• The D operator can be understood as shifting arbitrary predicates to a distributive interpretation with granularity *Atom* (i.e. singular individual):

(10)   **Definition: Atomic event-based D operator**

$$[\![D_\theta]\!] \stackrel{\text{def}}{=} \lambda P_{\langle v,t \rangle} \lambda e \left[e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right) \right]$$

(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose thematic roles θ are atoms.)

• Example:

(11)   The girls are wearing a dress.
       $\exists e [{}^*\text{agent}(e) = \bigoplus \text{girl} \wedge {}^*\text{wear}(e) \wedge \text{dress}(\text{theme}(e))]$
       (There is a potentially plural wearing event whose agents sum up to the girls, and whose theme is a dress.)

(12)   The girls D(are wearing a dress.)
       $\exists e [{}^*\text{agent}(e) = \bigoplus \text{girl} \wedge$
       $e \in {}^*\lambda e'(\text{wear}(e') \wedge \text{dress}(\text{theme}(e')) \wedge \text{Atom}(\text{agent}(e')))]$
       (There is an event whose agents sum up to the girls, and this event consists of wearing events for each of which the agent is a atom and the theme is a dress.)

- The star operator $^*\lambda e'$ is introduced through the D operator and takes scope over the predicate *dress* introduced by the theme.

Exercise 4.1 Which background assumptions ensure that (12) entails that each girl wears a dress? □

4.2.2   The leakage problem

- There are various other proposals on how to reformulate the D operator.

- Lasersohn 1998 proposes the following entry (among others):

(13)   **Distributivity operator over events (Lasersohn)**
$$[\![D^{Lasersohn}]\!] = \lambda P_{\langle e,vt\rangle}\lambda x\lambda e\forall y[y \leq_{Atom} x \rightarrow \exists e'[e' \leq e \land P(y)(e')]]$$

- This applies to a predicate of type $\langle e, vt\rangle$, e.g. $[\![\text{smile}]\!] = \lambda x\lambda e[\text{smile}(e) \land \text{agent}(e) = x]$.

- Inserting a D operator into *The girls smiled* before existential closure applies:

(14)   a.   Lasersohn's representation:
$\lambda e\forall y[y \leq_{Atom} \oplus \text{girl} \land \quad \rightarrow \exists e'[e' \leq e \land \text{smile}(e') \land {}^*\text{agent}(e') = y]$
     b.   My representation:
$\lambda e[{}^*\text{agent}(e) = \oplus \text{girl} \land e \in {}^*\lambda e'[\text{smile}(e') \land \text{Atom}({}^*\text{agent}(e'))]]$

- (14a) applies to all events that contain a smiling subevent for each girl, even if they also contain extraneous material. It suffers from what Bayer 1997 calls leakage. Whenever it (14a) applies to an event e, it also applies to any event of which e is a part.

- (14b) applies to all events that contain a smiling subevent for each girl and nothing else.

- Leakage causes problems in connection with event predicates such as *surprisingly*, *unharmoniously* or *in slow procession*.

- These predicates do not have divisive reference: they can hold of an event even if they do not hold of its parts (Schein 1993).

  (15)   Unharmoniously, every organ student sustained a note on the Wurlitzer.

- This says that the ensemble event was unharmonious and not any one student's note.

- Let *Lasersohn* stand for Lasersohn's (14a) and let *Mine* stand for my (14b).

- Imagine an event G that satisfies both *Lasersohn* and *Mine*, that is, the girls smiled in it.

- Let B be an event in which the boys cry.

- Now G ⊕ B does not satisfy *Mine*, but it does satisfy *Lasersohn*.

- Suppose that G is not surprising by itself, but that G ⊕ B is surprising. Then we have these judgments:

  (16)   a.   The girls smiled.   *true*
         b.   The girls smiled and the boys cried.   *true*
         c.   Surprisingly, the girls smiled.   **false**
         d.   Surprisingly, the girls smiled and the boys cried.   *true*

- If one of the D operators is applied to *smile*, then (17) is translated as (17a) or (17b).

  (17)   Surprisingly, the girls smiled.
         a. $\exists e[\text{surprising}(e) \wedge Lasersohn(e)]$
         b. $\exists e[\text{surprising}(e) \wedge Mine(e)]$

- The problem is that G ⊕ B satisfies both *Lasersohn* (by leakage) and the predicate *surprising* (by assumption). So Lasersohn's D operator wrongly predicts that (17) is judged true.

- The above implementation avoids this kind of leakage.

4.3   Atomic and nonatomic distributivity

- So far we have implemented the view called *atomic distributivity*: the D operator distributes over atoms, that is, over singular individuals (Lasersohn 1998, 1995, Link 1997, Winter 2001)

- Nonatomic view: phrasal distributivity may also quantify over nonatomic parts (Gillon 1987, 1990, Verkuyl & van der Does 1991, van der Does & Verkuyl 1996, Schwarzschild 1996, Brisson 1998, 2003, Malamud 2006a,b)

- Traditional argument is based on sentences like this, adapted from Gillon 1987:

  (18)   a.   Rodgers, Hammerstein, and Hart wrote *Oklahoma* and *On Your Toes.*
          b.   Rodgers, Hammerstein, and Hart wrote musicals.

- Rodgers, Hammerstein and Hart never wrote any musical together, nor did any of them ever write one all by himself. But Rodgers and Hammerstein wrote musicals like *Oklahoma* together, and Rodgers and Hart wrote musicals like *On your toes* together.

- On the basis of these facts, (18a) and (18b) are judged as true in the actual world, although it is neither true on the collective interpretation nor on an "atomically distributive" interpretation.

- The traditional nonatomic argument: in order to generate the reading on which (18b) is true, the predicates *wrote musicals* and *wrote Oklahoma and On Your Toes* must be interpreted as applying to nonatomic parts of the sum individual to which the subject refers.

- Generally implemented with covers (Gillon 1987): partitions of a set (19) or sum (20) whose cells/parts can overlap

(19) **Definition: Cover (set-theoretic)**
$$\text{Cov}(C, P) \stackrel{\text{def}}{=} \bigcup C = P \ \wedge \ \varnothing \notin C$$
(C is a cover of a set P if and only if C is a set of subsets of P whose union is P.)

(20) **Definition: Cover (mereological)**
$$\text{Cov}(C, x) \stackrel{\text{def}}{=} \bigoplus C = x$$
(C is a cover of a mereological object x is a set of parts of x whose sum is x.)

- Cover-based approaches modify the D operator to quantify over nonatomic parts of a cover of the plural individual.

- The first cover-based approaches assumed that the cover can be existentially quantified by the operator that introduces it:

(21) **Nonatomic distributivity operator, existentially bound cover**
$$[\![D_\exists]\!] = \lambda P_{\langle e,t \rangle} \lambda x \exists C[\text{Cov}(C, x) \ \wedge \ \forall y[C(y) \ \wedge \ y \le x \to P(y)]]$$

- On this view, sentences (18a) and (18b) are translated as follows:

(22) $\exists C[\text{Cov}(C, rodgers \oplus hammerstein \oplus hart) \ \wedge$
$\forall y[C(y) \ \wedge \ y \le x \to y \in [\![\text{wrote } Oklahoma \text{ and } On \ Your \ Toes]\!]]]$

(23) $\exists C[\text{Cov}(C, rodgers \oplus hammerstein \oplus hart) \ \wedge$
$\forall y[C(y) \ \wedge \ y \le x \to y \in [\![\text{wrote musicals}]\!]]]$

Exercise 4.2 For which value of C are these formulas true in the actual world? □

- Existentially bound covers are now generally considered untenable because they overgenerate nonatomically distributive readings

- Lasersohn 1989's problem: Suppose John, Mary, and Bill are the teaching assistants and each of them was paid exactly $7,000 last year. (24a) and (24b) are true, but (24c) is false.

  (24) a. *True:* The TAs were paid exactly $7,000 last year. *distributive*
        b. *True:* The TAs were paid exactly $21,000 last year. *collective*
        c. *False:* The TAs were paid exactly $14,000 last year. **nonatomically distributive*

- Giving up the existential cover-based operator $D_\exists$ in (21) explains why (24c) is false, because without this operator, there is no way to generate a true reading for this sentence.

- But now why are (18a) and (18b) true?

- As it turns out, the lexical cumulativity assumption is already enough (Lasersohn 1989):

  (25) $\forall w, x, y, z[\text{write}(w, x) \ \wedge \ \text{write}(y, z) \rightarrow \text{write}(w \oplus y, x \oplus z)]$

Exercise 4.3 What does this assumption translate to in a Neo-Davidsonian framework? □

- Further support: (28) is false in the actual world (Link 1997):

  (28)   Rodgers, Hammerstein and Hart wrote a musical.

  a. **True** if the three of them wrote a musical together – not the case. ✓ *collective*
  b. **True** if each of them wrote a musical by himself – not the case. ✓ *atomically distributive*
  c. **False** even though Rodgers and Hammerstein wrote a musical together, and Rodgers and Hart wrote another musical together. **nonatomically distributive*

- The absence of the nonatomically distributive reading of (28) is predicted if we give up the existential cover-based operator $D_\exists$.

- Lexical cumulativity derives the (available) nonatomically distributive reading of (18a) and (18b) but not the (unavailable) nonatomically distributive reading of (28):

(29) ⟦Rodgers, Hammerstein and Hart wrote *Oklahoma* and *On Your Toes*.⟧
= $\exists e[^*\text{write}(e) \wedge {}^*\text{agent}(e) = \text{rodgers} \oplus \text{hammerstein} \oplus \text{hart} \wedge {}^*\text{theme}(e) = \text{okl} \oplus \text{oyt}]$
(Allows for several writing events and for teamwork, as long as these two musicals are written.)

(30) ⟦Rodgers, Hammerstein and Hart wrote musicals.⟧
= $\exists e[^*\text{write}(e) \wedge {}^*\text{agent}(e) = \text{rodgers} \oplus \text{hammerstein} \oplus \text{hart} \wedge {}^*\text{musical}(^*\text{theme}(e))]$
(Allows for several writing events and for teamwork, and there can be several musicals in total.)

(31) ⟦Rodgers, Hammerstein and Hart wrote a musical.⟧
= $\exists e[^*\text{write}(e) \wedge {}^*\text{agent}(e) = \text{rodgers} \oplus \text{hammerstein} \oplus \text{hart} \wedge \text{musical}(^*\text{theme}(e))]$
(Allows for several writing events and for teamwork, but there has to be only one musical in total.)

- Lasersohn, as well as Winter 2001 and others, conclude from this and similar examples that the atomic approach to phrasal distributivity is superior to covers.

- However, Gillon 1990 and Schwarzschild 1996 identify a residue of cases in which a cover-based operator does seem necessary.

(32) **Scenario** Two pairs of shoes are on display, each pair with a \$50 price tag.
  a. The shoes cost \$100.   ✓ *collective (together)*
  b. The shoes cost \$25.   ? *atomically distributive (per shoe)*
  c. The shoes cost \$50. (Lasersohn 1995)   ✓ *nonatomically distributive (per pair)*

- Evidence that individual shoes, and not shoe pairs, are atoms in this context:

    (33)   How many shoes are on display? – Four / #Two.

    (34)   **Scenario** (Justin Bledin, p.c.) A magic square is a type of puzzle that involves filling in a grid with natural numbers so that the rows, columns, and diagonals all sum to the same thing. Suppose I give you a magic square to solve and I give you the following clue:

    (35)   The numbers sum to twenty-five.

- The salient reading of this sentence is not atomically distributive because it does not assert that each number sums to twenty-five, and it is not collective because it is each row, column and diagonal that sums to twenty-five, rather than the totality of the numbers.

- Schwarzschild 1996 proposes that the cover of D is anaphoric on context, and renames it Part:

    (36)   **Schwarzschild's nonatomic distributivity operator, free cover**
    $$[\![\text{Part}_C]\!] = \lambda P_{\langle e,t\rangle}\lambda x \forall y[C(y) \ \wedge \ y \le x \rightarrow P(y)]$$

- See Malamud 2012 for a decision-theoretic elaboration of this proposal.

- Nonatomic distributivity is always available for verbs, but for verb phrases it only occurs when context supplies a pragmatically salient cover. Atomic distributivity is available in both cases.

- To model nonatomic distributivity, I change the event-based atomic D operator repeated below as (37), by replacing the predicate *Atom* by a free predicate *C*. It plays the same role as the *C* predicate in Schwarzschild's operator:

    (37)   **Definition: Event-based D operator**
    $$[\![D_\theta]\!] \overset{\text{def}}{=} \lambda V \lambda e. \ e \in {}^*\lambda e' \left(\begin{array}{l} V(e') \ \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

Figure 4.1: V level versus verb phrase level distributivity in atomic domains

	lexical (V level)	phrasal (VP level)		lexical (V level)	phrasal (VP level)
atomic	available	available		meaning post.	Atomic D op.
nonatomic	available	only w. context		meaning post.	Cover-based Part op.

(a) Empirical generalization **(b)** Explanation

(Takes an event predicate *V* and returns a predicate that holds of any event *e* which can be divided into events that are in *V* and whose θs are atomic.)

(38) **Definition: Event-based Part operator**

$$[\![\text{Part}_{\theta,C}]\!] = \lambda V \lambda e.\ e \in {}^*\lambda e' \left(\begin{array}{c} V(e')\ \wedge \\ C(\theta(e')) \end{array} \right)$$

(Takes an event predicate *V* and returns a predicate that holds of any event *e* which can be divided into events that are in *V* and whose θs satisfy the contextually salient predicate *C*.)

- The Part operator has two parameters: dimension (thematic role) and granularity (C).

- Unlike Schwarzschild, I do not rely on pragmatics to ensure that *C* actually covers the θs of the event to which the output of the operator is applied.

- Following Schwarzschild 1996, I assume that the *C* parameter of the Part operator in (38) can only be set in one of two ways: either it is set to the predicate *Atom* or to an anaphorically salient level of granularity.

4.4 The scopal behavior of *for*-adverbials

- Assume as a baseline analysis that *for* is a universal quantifier over instants, as if it was the temporal counterpart of *every*:

(39)   ⟦for an hour⟧ (baseline) =
$\lambda P_{\langle v,t \rangle} \exists t[\text{hours}(t) = 1 \wedge \forall t'[t' <_{Atom} t \rightarrow \exists e[P(e) \wedge \tau(e) = t')]]$

- Given this assumption, the scopal behavior of *for*-adverbials is surprising (Carlson 1977, Zucchi & White 2001):

(40) a. John pushed a cart for an hour. $\exists > \forall$; *$\forall > \exists$
      b. I dialed a wrong phone number for five minutes. $\exists > \forall$; *$\forall > \exists$
      c. She bounced a ball for 20 minutes. $\exists > \forall$; *$\forall > \exists$
      d. He kicked a wall for a couple of hours. $\exists > \forall$; *$\forall > \exists$
      e. She opened and closed a drawer for half an hour. $\exists > \forall$; *$\forall > \exists$
      f. I petted a rabbit for two hours. $\exists > \forall$; *$\forall > \exists$

- Same pattern in German (Kratzer 2008). This is remarkable since quantifier scope in German normally follows surface order.

(41) a. Ich hab' fünf Minuten lang eine falsche Telefonnummer
          I   have five minutes long a   wrong telephone.number
          gewählt.
          dialed.
      b. Ich hab' eine falsche Telefonnummer   fünf Minuten lang
          I   have a   wrong telephone.number five minutes long
          gewählt.
          dialed.

- Same pattern even when the missing reading would be more plausible:

(42) ??John found a flea on his dog for a month. (Zucchi & White 2001)

- Plural indefinites pattern like singular indefinites:

(43)   John saw thirty zebras for three hours. $30 > \forall$; *$\forall > 30$

- Bare plurals and bare mass nouns are an exception (Verkuyl 1972):

(44)   a.   John discovered fleas on his dog for six weeks.
       b.   John discovered crabgrass in his yard for six weeks.

(45)   a.   Tourists discovered that quaint little village for years.
       b.   Water leaked through John's ceiling for six months.

- Another exception is the quantifier *no* (Rooth 1995)

(46)   John ate no cookies for thirty days.

- Overt temporal adverbials, call them *interveners*, can "trap" the indefinite:

  – Context-dependent temporal definites (Deo & Piñango 2011)

  (47)   Jane ate an egg/two eggs **at breakfast** for a month.

  – Temporal universal quantifiers (Zucchi & White 2001)

  (48)   John found a flea on his dog **every day** for a month.

  – Pluractional adverbials (see also Beck & von Stechow 2007)

  (49)   John found a flea on his dog **day after day** for a month.

- Compare the baseline analysis with the one in (51):

(50)   ⟦for an hour⟧ (baseline)
       $= \lambda P_{\langle v,t\rangle} \exists t[\mathrm{hours}(t) = 1 \wedge \forall t'[t' <_{Atom} t \rightarrow \exists e[P(e) \wedge \tau(e) = t')]]]$

(51)   ⟦for an hour⟧ (my proposal)
       $= \lambda P_{\langle v,t\rangle} \lambda e : \mathrm{SR}_{\tau,\lambda t'.t'<\tau(e)}(P)(e).\ P(e) \wedge \mathrm{hours}(\tau(e)) = 1$

- This immediately predicts that the indefinite in (40a) must take wide scope.

(52)   John pushed a cart for an hour.   = (40a)

(53)   ⟦push a cart⟧ $= \lambda e[{}^{*}\mathrm{push}(e) \wedge \mathrm{cart}({}^{*}\mathrm{theme}(e))]$
       (True of any pushing event or sum of pushing events whose theme
       is one and the same cart.)

(54)   ⟦John pushed a cart for an hour⟧
= $\exists e : \mathrm{SR}_{\tau, \lambda t'. t' < \tau(e)} (\lambda e' [^* \mathrm{push}(e') \wedge \mathrm{cart}(^* \mathrm{theme}(e'))])(e).$
$[^* \mathrm{agent}(e) = j \wedge {}^* \mathrm{push}(e) \wedge \mathrm{cart}(^* \mathrm{theme}(e)) \wedge \mathrm{hours}(\tau(e)) = 1]$
(There is a pushing event or sum of pushing events whose theme is one cart, whose agent is John, and whose runtime measures one hour.)

- As Kratzer 2008 points out, lexical cumulativity extends this style of account even to achievement verbs like *find*.

(55)   ⟦John found a flea for a month⟧
= $\exists e : \mathrm{SR}_{\tau, \lambda t'. t' < \tau(e)} (\lambda e' [^* \mathrm{find}(e') \wedge \mathrm{flea}(^* \mathrm{theme}(e'))])(e).$
$[^* \mathrm{agent}(e) = j \wedge {}^* \mathrm{find}(e) \wedge \mathrm{flea}(^* \mathrm{theme}(e)) \wedge \mathrm{months}(\tau(e)) = 1]$
(There is a finding event or sum of finding events whose theme is one flea, whose agent is John, and whose runtime measures one month.)

- Plurals and mass terms behave differently because they have cumulative reference.

(56)   ⟦John found fleas for a month⟧
= $\exists e : \mathrm{SR}_{\tau, \lambda t'. t' < \tau(e)} ([^* \mathrm{find}(e') \wedge {}^* \mathrm{flea}(^* \mathrm{theme}(e'))])(e).$
$[^* \mathrm{agent}(e) = j \wedge {}^* \mathrm{find}(e) \wedge {}^* \mathrm{flea}(^* \mathrm{theme}(e)) \wedge \mathrm{months}(\tau(e)) = 1]$
(There is a finding event or sum of finding events whose theme is a set of fleas, whose agent is John, and whose runtime measures one month.)

- This presupposition is fulfilled in this case (why?).

4.5   Nonatomic distributivity in *for*-adverbials

- We have seen that indefinites can't covary with the *for*-adverbial.

(57)   ??John found two fleas on his dog for a month.

(58)   a.   ??John noticed a discrepancy/two discrepancies for a week.
         b.   ??John discovered a new proof/two new proofs for a week. (Deo & Piñango 2011)

- But several authors have noted that there are exceptions.

(59)   **Context:** the patient's daily intake is discussed.
         The patient took two pills for a month and then went back to one pill.
         (Moltmann 1991)

(60)   **Context:** the bicycle is designed to carry around three children at a time.
         This bicycle carried three children around Amsterdam for twenty years.
         (Rothstein 2004)

(61)   We built a huge snowman in our front yard for several years.
         (Deo & Piñango 2011)

- This inference takes time (self-paced reading tests, Todorova, Straub, Badecker & Frank 2000):

(62)   a.   Even though Howard sent a large check to his daughter for many years, she refused to accept his money. *longer reading time*

         b.   Even though Howard sent large checks to his daughter for many years, she refused to accept his money. *shorter reading time*

- This looks like nonatomic phrasal distributivity!

- Nonatomic lexical distributivity is shown in examples like this:

(63)   Five thousand people gathered near Amsterdam. (van der Does 1993)

- Here the predicate *gather near Amsterdam* can be applied distributively (i.e. several gatherings) to nonatomic entities (a single person cannot gather)

- Nonatomic phrasal distributivity is usually unavailable (e.g. Lasersohn 1989):

  (64)   Rogers, Hammerstein and Hart wrote a musical.

  a.   **True** if the three of them wrote a musical together.   ✓ *collective*
  b.   **True** if each of them wrote a musical by himself.   ✓ *atomically distributive*
  c.   **False** if Rodgers and Hammerstein wrote a musical together, and Rodgers and Hart wrote another musical together. **nonatomically distributive*

  (65)   **Scenario** Al, Bill, Jim and Ed each weigh 100kg.
  a.   **True**: The men weigh 300kg.   ✓ *collective (together)*
  b.   **True**: The men weigh 100kg.   ✓ *atomically distributive (per man)*
  c.   **False**: The men weigh 200kg. **nonatomically distributive (per pair)*

- Exception: a level of granularity is made salient through context or world knowledge (Lasersohn 1995, Schwarzschild 1991, 1996)

  (66)   **Scenario** Two pairs of shoes are on display, each pair with a $50 price tag.
  a.   The shoes cost $100.   ✓ *collective (together)*
  b.   The shoes cost $25.   *? atomically distributive (per shoe)*
  c.   The shoes cost $50.   ✓ *nonatomically distributive (per pair)*

- Nonatomic distributivity is much easier with a bare plural (Link 1997):

  (67)   a.   Rodgers, Hammerstein and Hart wrote a musical. **pairwise*
          b.   Rodgers, Hammerstein, and Hart wrote musicals.   *pairwise*

Table 4.1: V level versus verb phrase level distributivity in atomic domains

	lexical (V level)	phrasal (VP level)	lexical (V level)	phrasal (VP level)
atomic	available	available	meaning post.	Atomic D op.
nonatomic	available	only w. context	meaning post.	Cover-based Part op.
	(a) Empirical generalization		**(b)** Explanation	

Table 4.2: Distributivity in the temporal domain

	lexical (V level)	phrasal (VP level)	lexical (V level)	phrasal (VP level)
atomic	**n/a**	**n/a**	**n/a**	**n/a**
nonatomic	available	only w. context	lexical cum.	Cover-based Part op.
	(a) Empirical generalization		**(b)** Explanation	

- Following Schwarzschild 1996, we can model the context dependency of nonatomic distributivity by assuming that there is a VP-level Part operator that contains an anaphoric cover over contextually salient entities (pairs of shoes, etc.).

- In a domain like time, there are no atomic covers. So setting dimension to τ should be incompatible with setting granularity to *Atom*.

(68) ??John found two fleas on his dog for a month.

(69)   The patient took two pills for a month.

- The predicate *find two fleas* by itself does not satisfy the presupposition of the *for*-adverbial (why not?). Same for *take two pills*.

- The following theorem leads us to expect that the Part operator should function as a repair strategy whenever a *for*-adverbial imposes a stratified reference presupposition. More specifically, we expect this repair strategy

to succeed if and only if the salient level of granularity is at least as coarse as the granularity parameter of the *for*-adverbial.

(70)   **Theorem: Part$_{\theta,C}$ leads to universal stratified reference**
$\forall V \forall \theta \forall C \forall C'[C \subseteq C' \rightarrow \mathrm{SR}_{\theta,C'}(\mathrm{Part}_{\theta,C}(V))]$
(When the Part operator, coindexed with thematic role θ and with granularity threshold C, is applied to any predicate, the result has universal stratified reference with respect to θ and any C' that is at least as coarse as C.)

- See proof in the Appendix of Champollion 2017.

- Assume $\lambda t[\mathrm{days}(t) \leq 1]$ ("once a day") is salient.

- It follows that (69) can be interpreted by applying $\mathrm{Part}_{\tau,\mathrm{days}(t)\leq 1}$ to its verb phrase:

(71) $\mathrm{Part}_{\tau,\mathrm{days}(t)\leq 1}(\lambda e[\,{}^*\mathrm{take}(e) \wedge {}^*\mathrm{pill}({}^*\mathrm{theme}(e)) \wedge |{}^*\mathrm{theme}(e)| = 2])$
$= \lambda e[e \in {}^*\lambda e'\,({}^*\mathrm{take}(e') \wedge {}^*\mathrm{pill}({}^*\mathrm{theme}(e')) \wedge |{}^*\mathrm{theme}(e')| = 2] \wedge \mathrm{days}(\tau(e$
(True of any plural event that consists of one or more events of taking two pills which each take place within a day.)

(72) $\exists e[\,{}^*\mathrm{agent}(e) = \mathrm{the.patient} \wedge \mathrm{months}(\tau(e)) = 1 \wedge$
$e \in {}^*\lambda e'\,({}^*\mathrm{take}(e') \wedge {}^*\mathrm{pill}({}^*\mathrm{theme}(e')) \wedge |{}^*\mathrm{theme}(e')| = 2] \wedge \mathrm{days}(\tau(e')) \leq$
(There is a plural event that consists of one or more events of taking two pills which each take place within a day. Its agent is the patient, and its runtime measures a month.)

- For plural indefinites, there is an extra wrinkle:

(73)   John saw thirty zebras for three hours. $\checkmark 30 > \forall; \, {}^*\forall > 30$

- Predicates like *see thirty zebras* do not have the subinterval property: picture a safari in which the thirty zebras are seen in succession.

(74) $[\![\text{see thirty zebras}]\!] = \lambda e[\,{}^*\mathrm{see}(e) \wedge {}^*\mathrm{zebra}({}^*\mathrm{theme}(e)) \wedge |{}^*\mathrm{theme}(e)| = 30]$

(True of any possibly plural event in which a total of thirty zebras are seen)

- However, applying distributive QR (called SQI in Landman 1996) to *thirty zebras* leaves a trace behind whose value is an atomic individual x. For any fixed atomic x, the predicate *see x* has the subinterval property. In this case, QR is driven by the need to satisfy the presupposition of the *for*-adverbial.

(75) $[\![\text{see t}_1]\!] = \lambda e [\,^*\text{see}(e) \wedge {}^*\text{theme}(e) = g(1)]$
(True of any possibly plural event in which the atomic individual $g(1)$ is seen)

- One can then nondistributively QR *three hours* above *thirty zebras* to ensure that the three-hour timespan does not covary with the zebras.

(76) [[three hours] $\lambda 2$ [[thirty zebras (each)] $\lambda 1$ [[John saw t$_1$] [for t$_2$]]]]

- In the LF in (76), the silent *each* stands for the effect of distributive QR.

4.6 Modeling interveners

- We have seen that temporal definites, quantifiers, and pluractionals can lead to different-object effects:

(77) a. Jane ate an egg **at breakfast** for a month.
      b. John found a flea on his dog **every day** for a month.
      c. John found a flea on his dog **day after day** for a month.

- *Every day* and *day by day* have the same entry as the Part operator, except that the granularity is hard-wired to "day" instead of being given by context:

(78) $[\![\text{every day}]\!] = [\![\text{day by day}]\!] \overset{\text{def}}{=} \lambda P_{\langle v,t \rangle} \lambda e [e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \mathbf{days}(\tau(e')) \le \mathbf{1} \end{array} \right)]$

(Takes an event predicate *P* and returns a predicate that holds of any event *e* which consists entirely of events that are in *P* and whose runtimes **are at most a day long**)

- This leads to an interpretation like the following (ignoring the presupposition of *for a month* for convenience):

(79)   a.   ⟦John [[[found a flea] every day] for a month]⟧

        b. $\exists e[{}^{*}\text{agent}(e) = \text{john} \;\wedge\; \text{months}(\tau(e)) = 1 \;\wedge$
$e \in {}^{*}\lambda e'\,({}^{*}\text{find}(e') \;\wedge\; {}^{*}\text{flea}({}^{*}\text{theme}(e')) \;\wedge\; |{}^{*}\text{theme}(e')| = 1] \;\wedge\; \text{days}($
(There is a plural event that consists of one or more events of finding a flea which each take place within a day. Its agent is John, and its runtime measures a month.)

- *At breakfast* is treated similarly, except that granularity is hard-wired to "breakfast".

(80)   ⟦at breakfast⟧ $\overset{\text{def}}{=} \lambda P_{\langle v,t\rangle}\lambda e[e \in {}^{*}\lambda e'\,(P(e') \;\wedge\; \exists \mathbf{i}[\mathbf{breakfast}(\mathbf{i}) \wedge \tau(e') \leq \mathbf{i}]))]$
(Takes an event predicate *P* and returns a predicate that holds of any event *e* which consists entirely of events that are in *P* and whose runtimes **are within a breakfast**)

- These entries should be integrated into a more extensive account of temporal dependencies, see Pratt & Francez 2001, von Stechow 2002 and Champollion 2011. This is not trivial.

5

Overt distributivity

Building on Champollion 2016b, this lecture explains the crosslinguistic semantic differences between distance-distributive items such as English each *and German* jeweils *by treating them as overt versions of the atomic distributivity operator D and the nonatomic distributivity operator Part respectively. The proposed analysis in terms of stratified reference explains why* jeweils *can distribute over salient occasions and why this is never possible for* each *(Zimmermann 2002). It also accounts for the fact that distributive determiners can take part in cumulative readings with items outside of their syntactic scope, and for their ability to interact with nondistributive event modifiers (Schein 1993, Champollion 2010a).*

5.1   Introduction

- There are three uses of 'each' in English:

  (1)   a.   **Adnominal each:** Two men have carried three suitcases *each*.
        b.   **Adverbial each:** Two men have *each* carried three suitcases.
        c.   **Determiner each:** *Each* man has carried three suitcases.

- This lecture focuses on data and observations by Zimmermann 2002.

- In German, adnominal and adverbial distance-distributive (DD) items are different from the distributive determiner:

  (2)   a.   **Adnominal:** Die Männer haben *jeweils* drei Koffer getragen.
        b.   **Adverbial:** Die Männer haben *jeweils* drei Koffer getragen.
        c.   **Determiner:** *Jeder/*Jeweils* Mann hat drei Koffer getragen.

- Though adverbial and adnominal *jeweils* are similar, they can be teased apart syntactically. For more details, see Zimmermann.

- *Each* and *jeweils* generalize to two classes of DD items:

  – *Each*-type DD items can also be used as determiners.

  – *Jeweils*-type DD items cannot double as determiners.

- **Zimmermann's Generalization** (illustrated below): All *each*-type DD items can only distribute over individuals. This contrasts with many *jeweils*-type DD items, which can also distribute over occasions (= salient chunks of time or space).

5.1.1   Questions.

– How can we capture the synonymy of the determiner, adnominal and adverbial uses of *each* in English?

– How can we represent the fact that DD items across languages share some part of their meanings?

– How do DD items fit into distributivity theory more generally? How can we formally capture the semantic variation among DD items?

– How can we explain Zimmermann's Generalization?

5.2   Illustrating Zimmermann's Generalization

5.2.1   *Jeweils*-type DD Items

Occur in Bulgarian, Czech, German, Korean, Polish, Romanian, and Russian (Zimmermann a.o.)

- These languages have adnominal DD items that cannot double as determiners.

- All of these DD items can distribute over individuals, like English *each*.

- All of them can also (given supporting context) distribute over temporal/spatial intervals or 'occasions':

> (3)   Die Kinder   haben jeweils zwei Affen   gesehen.
>   The children have   DIST   two monkeys seen.
> > *German*
>
> > a.   *Always available:* 'Each of the children has seen two monkeys'
> > b.   *Available, though only with supporting context:* 'The children have seen two monkeys on each occasion'

> (4)   Hans hat jeweils zwei Affen   gesehen.
>   Hans has DIST   two monkeys seen.
> > *German*
>
> 'Hans has seen two monkeys on each occasion'

5.2.2   *Each*-type DD Items

> Occur in Albanian, Dutch, English, French, Icelandic, Italian, Japanese, Norwegian, Portuguese, Russian, and possibly Latin (Zimmermann a.o.)

- These languages have adnominal DD items that can also be used as distributive determiners.

- All of these DD items can distribute only over individuals, not over occasions:

> (5)   The children have seen two monkeys each.
> > a.   *Available:* 'Each of the children has seen two monkeys'

b.   *Unavailable:* 'The children have seen two monkeys on each occasion'

(6)   *John has seen two monkeys each.
      Intended: John has seen two monkeys on each occasion.

5.2.3   A Note on Reduplication: A Note on Reduplication

- Many languages express adnominal distance distributivity by reduplicating a numeral (Gil 1982). In this category, we both find cases where reduplication does not give rise to occasion readings, such as Hungarian (Farkas 1997, Szabolcsi 2010), and cases where it does, such as Telugu (Balusu 2005); also Hausa (Zimmermann 2008), Karitiana (Müller & Negrão 2012) and Telugu (Balusu 2005, Balusu & Jayaseelan 2013) and the suffix in Tlingit (Cable 2014).

(7)   pilla-lu1 renDu renDu kootu-lu-ni   cuus-ee-ru.
      kid-Pl   two   two   monkey-Pl-Acc see-Past-3PPl.
                              Telugu

      a.   'Each of the children has seen two monkeys'
      b.   'The children have seen two monkeys on each occasion'

(8)   Raamu renDu renDu kootu-lu-ni   cuus-ee-Du.
      Ram   two   two   monkey-Pl-Acc see-Past-3PSg.
                              Telugu

      'Ram has seen two monkeys on each occasion'

- The import of these cases on Zimmermann's generalization is unclear, as reduplication is not expected to be able to act as a determiner.

5.3   Capturing the Semantic Variation

- I propose that adnominal *each* and *jeweils* include two versions of the distributivity operator (cf. Link 1998b for a similar claim for German *je*, a short form of *jeweils*).

- In Lecture 4, I have motivated a formulation of the D operator that equips it with a dimension parameter and a granularity parameter, just like stratified reference.

  (9)   **Definition: Atomic event-based D operator**
  $$[\![D_\theta]\!] \overset{\text{def}}{=} \lambda P_{\langle v,t\rangle}\lambda e\big[e \in {}^*\lambda e'\begin{pmatrix} P(e') \wedge \\ \text{Atom}(\theta(e')) \end{pmatrix}\big]$$
  (Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose thematic roles θ are atoms.)

  (10)   **Definition: Event-based Part operator with covers**
  $$[\![\text{Part}_{\theta,\mathbf{C}}]\!] \overset{\text{def}}{=} \lambda P_{\langle v,t\rangle}\lambda e\big[e \in {}^*\lambda e'\begin{pmatrix} P(e') \wedge \\ \mathbf{C}(\theta(e')) \end{pmatrix}\big]$$
  (Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs **satisfy the contextually salient 'cover predicate' C**.)

- The thematic role parameter allows us to capture the fact that DD items can also target different thematic roles (Zimmermann 2002):

  (11)   The boys told the girls two stories each.   *Target: agent*
              (two stories per boy)

  (12)   The boys told the girls two stories each.   *Target: recipient*
              (two stories per girl)

- When we instantiate C with *Atom*, we get Link's VP-level D operator.

- Schwarzschild's Part operator is a generalization: instead of specifying the granularity parameter to be atomic, we leave it free.

- We can use these parameters to describe the difference between *each* and *jeweils*.

    - *Each* includes the atomic distributivity operator of Link 1987, which can only distribute over count domains (granularity=atom).

    - *Jeweils* includes the cover-based distributivity operator of Schwarzschild 1996, which can also distribute over noncount domains like time (granularity=contextual).

- The setting "granularity=atom" blocks "dimension=time" because time is continuous and noncount – there are no atoms to distribute over.

- I propose that adnominal *each* comes prespecified for "granularity=atom". This blocks "dimension=time", so distributivity over occasions is unavailable. *Jeweils* does not come prespecified for anything.

5.4   Explaining Zimmermann's Generalization

- I propose that in English, adnominal, adverbial and *each* have identical meanings up to type-shifting.

- Determiner *each* is only compatible with count domains ("granularity=atomic") – **each mud, *each water* etc.

- Adnominal *each* is formally identical, so it inherits this property.

- The count domain restriction of adnominal *each* is incompatible with time because they do not contain atoms: "granularity=atom" blocks "dimension=time".

- *Jeweils*-type DD items are formally different from determiners. So it is unsurprising that they do not inherit "granularity=atom".

5.5   Formalization

- We use type shifters for composition like these:

  (13)   a.   Type shifter for indefinites: $\lambda\theta_{ve}\lambda P_{et}\lambda V_{vt}\lambda e[V(e)\wedge P(\theta(e))]$
          b.   Type shifter for definites: $\lambda\theta_{ve}\lambda x\lambda V_{vt}\lambda e[V(e)\wedge\theta(e)=x]$

- Each of these type shifters combines a noun phrase with its theta role head to build an event predicate modifier of type $\langle vt,vt\rangle$.

- Example:

  (14) $[\![[\text{agent [the boys]}]]\!] = \lambda V\lambda e[V(e)\wedge {}^*\text{agent}(e) = \bigoplus \text{boy}]$

  (15) $[\![[\text{theme [two monkeys]}]]\!] = \lambda V\lambda e[V(e)\wedge|{}^*\text{theme}(e)| = 2\wedge {}^*\text{monkey}({}^*\text{theme}(e))]$

- The event variable is existentially bound if the sentence is uttered out of the blue (see Figure 5.1). If the sentence is understood as referring to a specific event, the event variable is instead resolved to that event.

- Adverbial *each* is a VP modifier, and is synonymous to Link's D operator:

  (16) $[\![\text{each}_\theta]\!]_{adverbial} = [\![D_{\theta,Atom}]\!] = (10)$

- Adnominal and determiner *each* need to be type-shifted, but both are defined in terms of D:

  (17) $[\![\text{each}_\theta]\!]_{adnominal} = \lambda P_{et}\lambda\Theta_{ve}\lambda V_{vt}\lambda e\,[\![D_{\theta,Atom}]\!](\lambda e'[V(e')\wedge P(\Theta(e'))])(e)]$

- Adnominal *each* combines with an indefinite NP and then with a theta head:

  (18) $[\![[[[\text{two monkeys] each}_{ag}]\text{ theme}]]\!]$
         $= \lambda V_{vt}\lambda e[[\![D_{ag,Atom}]\!](\lambda e'[V(e')\wedge|{}^*\text{theme}(e')|=2\wedge {}^*\text{monkey}({}^*\text{theme}(e'))$

CP
$\exists e.$
$e \in [\![D_{agent,\,Atom}]\!](\lambda e'.{}^*\text{see}(e)$
$\wedge\ \text{two-monkeys}({}^*\text{theme}(e')))]$
$\wedge\ {}^*\text{agent}(e) = \bigoplus \text{boy}$

[closure]
$\lambda V \exists e. V(e)$

IP
$\lambda e.$
$e \in [\![D_{agent,\,Atom}]\!](\lambda e'.{}^*\text{see}(e)$
$\wedge\ \text{two-monkeys}({}^*\text{theme}(e')))]$
$\wedge\ {}^*\text{agent}(e) = \bigoplus \text{boy}$

DP
$\lambda V \lambda e. V(e)$
$\wedge\ {}^*\text{agent}(e) = \bigoplus \text{boy}$

The boys [agent]

VP
$\lambda e.$
$e \in [\![D_{agent,\,Atom}]\!](\lambda e'.{}^*\text{see}(e)$
$\wedge\ \text{two-monkeys}({}^*\text{theme}(e')))]$

saw
$\lambda e.{}^*\text{see}(e)$

DP
$\lambda V_{\langle v,t\rangle} \lambda e.$
$e \in [\![D_{agent,\,Atom}]\!](\lambda e'[V(e')\wedge$
$\text{two-monkeys}({}^*\text{theme}(e'))])$

[theme]
${}^*\text{theme}$

$\lambda \Theta_{\langle ve\rangle} \lambda V_{\langle v,t\rangle} \lambda e.$
$e \in [\![D_{agent,\,Atom}]\!](\lambda e'[V(e')\wedge$
$\text{two-monkeys}(\Theta(e'))])$

NP
$\lambda x.\text{two-monkeys}(x)$

two monkeys

each_{agent}
$\lambda P_{\langle e,t\rangle} \lambda \Theta_{\langle ve\rangle} \lambda V_{\langle v,t\rangle} \lambda e.$
$e \in [\![D_{agent,\,Atom}]\!](\lambda e'[V(e') \wedge P(\Theta(e'))])$
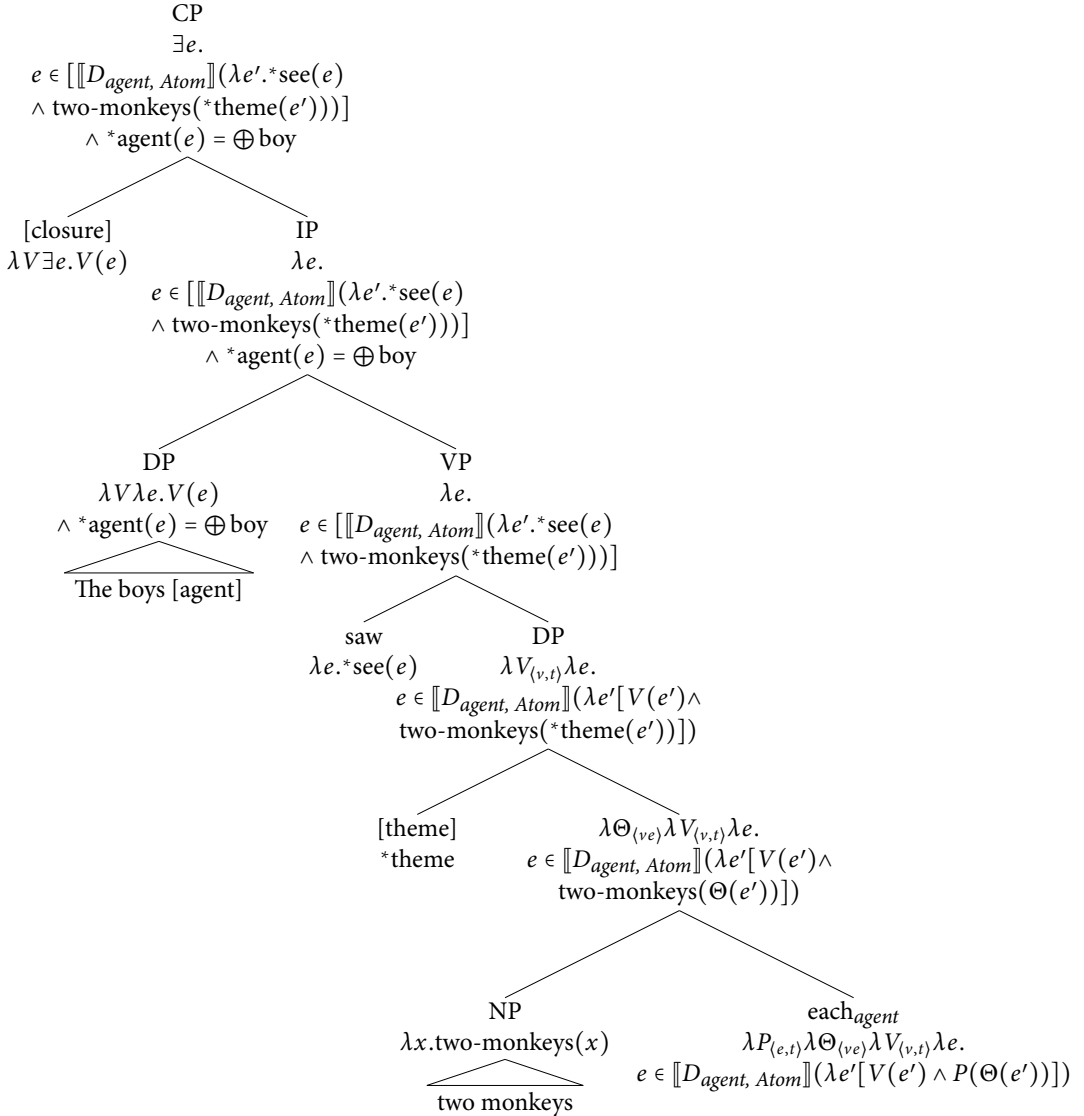
Figure 5.1: Deriving *The boys saw two monkeys each.*

$$= \lambda V_{vt}\lambda e[e \in {}^*\lambda e'[{}^*\text{see}(e') \wedge |{}^*\text{theme}(e')| = 2 \wedge {}^*\text{monkey}({}^*\text{theme}(e')) \wedge \text{Atom}(\text{agent}(e'))]]$$

(19) $\llbracket \text{each} \rrbracket_{determiner} = \lambda P_{et}\lambda \theta_{ve}\lambda V_{vt}\lambda e \; [\theta(e) = \bigoplus P \wedge \llbracket D_{\theta,Atom} \rrbracket (V)(e)]$

Determiner *each* combines first with a nominal and then with a theta head:

(20) $\llbracket [[\text{Each child] agent}] \rrbracket$
$= \lambda V_{vt}\lambda e[{}^*\text{agent}(e) = \bigoplus \text{child} \wedge \llbracket D_{ag,Atom} \rrbracket (V)(e)]$
$= \lambda V_{vt}\lambda e[{}^*\text{agent}(e) = \bigoplus \text{child} \wedge e \in {}^*\lambda e'[V(e') \wedge \text{Atom}(\text{agent}(e'))]]$

The result of these derivations is always the same, which reflects their synonymy:

(21) $\llbracket$The children each$_{ag}$ saw two monkeys$\rrbracket$
$= \llbracket$The children saw two monkeys each$_{ag}\rrbracket$
$= \llbracket$Each child saw two monkeys$\rrbracket$
$= \llbracket$The children D$_{ag,\;Atom}$ saw two monkeys$\rrbracket$

Adverbial *jeweils* is treated as in (22).

(22) $\llbracket \text{jeweils}_{\theta,C} \rrbracket_{adverbial}$
$= \lambda V\lambda e[\llbracket \text{Part}_{\theta,C} \rrbracket (V)(e) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e))]$
$= \lambda V\lambda e[e \in {}^*\lambda e'\begin{pmatrix} V(e') \wedge \\ C(\theta(e')) \end{pmatrix} \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e))]$

- Essentially the same type shift as in (17) brings us to adnominal *jeweils*:

(23) $\llbracket \text{jeweils}_{\theta,C} \rrbracket_{adnominal}$
$= \lambda P\lambda \Theta \lambda e[\llbracket \text{Part}_{\theta,C} \rrbracket (\lambda e'[P(\Theta(e'))])(e) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e))]$
$= \lambda P\lambda \Theta \lambda e[e \in {}^*\lambda e'\begin{pmatrix} P(\Theta(e')) \wedge \\ C(\theta(e')) \end{pmatrix} \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e))]$

- Setting C to *Atom* and θ to *agent* leads to distribution over individuals:

(24)   Die Kinder   haben jeweils$_{ag,Atom}$ zwei Affen   gesehen.
        The children have   DIST   two monkeys seen.
        "The children have each seen two monkeys."

- With supporting context, the anaphoric predicate C can be set to a salient antecedent other than *Atom*. Then θ is free to adopt values like τ (runtime).

- Example: the children have been to the zoo to see animals last Monday, last Wednesday and last Friday. (25) is uttered with reference to that sum event:

(25)   ⟦Die Jungen haben jeweils$_{\tau,zoovisit}$ zwei Affen gesehen.⟧ =
        $\exists e.\ {}^{*}\text{agent}(e) = \oplus\,\text{boy} \land {}^{*}\text{see}(e)$
        $\land\ e \in {}^{*}\lambda e' \left(\begin{array}{l} |{}^{*}\text{theme}(e')| = 2 \ \land\ {}^{*}\text{monkey}({}^{*}\text{theme}(e')) \land \\ \text{zoovisit}(\tau(e')) \end{array} \right)$
        $\land\ \oplus\,\text{zoovisit} = \tau(e)$
        'The boys have seen two monkeys on each salient occasion (that is, on each of the three zoo visits).'

- Since *zoovisit* is salient, C can be resolved to it rather than to *Atom*.

- Since there are no atoms in time, it is only now that θ can be set to τ, rather than to *agent* as in (24).

- What (25) asserts in this context is the following:

    - there is an event e whose (discontinuous) runtime is the sum of the three zoo visits; that this event has the boys as its agents

    - it can be divided into subevents, each of whose runtimes is the time of a zoo visit;

    - and each of these subevents is an event whose theme are two monkeys.

- That these subevents are seeing events is entailed by the fact that *see* is lexically distributive on its theme argument, which in turn can be formally represented as a meaning postulate, as discussed in Section 2.9.

- I assume that runtime is closed under sum just like other thematic roles ($\tau = {}^*\tau$), or in other words, it is a sum homomorphism. This means that any way of dividing e must result in parts whose runtimes sum up to $\tau(e)$.

- The conjunct $\bigoplus$ zoovisit $= \tau(e)$ makes sure that $\tau(e)$ is the sum of the times of the three zoo visits in question. Hence each of these zoo visits must be the runtime of one of the seeing-two-monkeys events.

- Without it, (25) would be predicted true even if the boys failed to see two monkeys on some of the salient zoo visits.

5.6   Summary

How can we capture the synonymy of the determiner, adnominal and adverbial uses of *each* in English?

- They are all derived from Link's D operator.

How can we represent the fact that DD items across languages share some part of their meanings?

- They are derived from related distributivity operators (Link's or Schwarzschild's) which differ only in their parameter settings.

How do DD items fit into distributivity theory more generally? How can we formally capture the semantic variation among DD items?

- They display the same parametric variation as other flavors of distributivity do.

How can we explain Zimmermann's generalization?

- *Each*-type DD items are formally identical to determiners and therefore inherit their "granularity=atomic" value. *Jeweils*-type DD items may have any setting for the granularity parameter.

5.7   Distributive determiners and cumulative readings

- Both *every* and *each* are distributive:

  (26)   #Every/#Each soldier surrounded the castle.   (Kroch 1974:
          Chapter 5)

- Traditionally, the determiners *every* and *each* are analyzed in terms of universal quantification (e.g., Montague 1973):

  (27) $[\![\text{every boy}]\!] = \lambda P \forall x [\text{boy}(x) \rightarrow P(x)]$ (traditional)

- I adopt a different analysis (see also Kratzer 2000, Ferreira 2005, and Thomas 2015) that makes the connection to the D operator prominent:

  (28) $[\![\text{each}]\!]_{\text{determiner}} = [\![\text{every}]\!]$
          $= \lambda P \lambda \theta \lambda V \lambda e.\ \theta(e) = \bigoplus P \wedge [\![D_\theta]\!](V)(e)$
          $= \lambda P \lambda \theta \lambda V \lambda e.\ \theta(e) = \bigoplus P \wedge e \in {}^*\lambda e' \left(\begin{array}{l} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$

- Sample noun phrase denotation and sentence:

  (29) $[\![[\text{agent}]\ \text{every boy}]\!]$
          $= \lambda V \lambda e.\ {}^*\text{agent}(e) = \bigoplus \text{boy} \wedge e \in {}^*\lambda e' \left(\begin{array}{l} V(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$
          (Takes an event predicate V and returns a predicate that holds of any event e whose agent is all the boys and which consists entirely of events that are in V and whose agents are individual boys.)

  (30) $[\![[\text{agent}]\ \text{every boy carried three suitcases}]\!]$
          $= \exists e.\ {}^*\text{agent}(e) = \bigoplus \text{boy} \wedge$
          $e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\text{carry}(e') \wedge |{}^*\text{theme}(e')| = 3 \wedge {}^*\text{suitcase}({}^*\text{theme}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$

(This says that there is an event *e* whose agent is all the boys and which consists entirely of carrying events whose agents are individual boys and whose themes are sums of three suitcases.)

- This helps with examples like the following (Taylor 1985, Schein 1993, Kratzer 2000, Ferreira 2005, Champollion 2010a, Thomas 2015):

(31) a. Unharmoniously, every organ student sustained a note on the Wurlitzer for sixteen measures.   (Schein 1993)
      b. In a complete lack of harmony, each monk started to sing the Kyrie in a different mode.   (Thomas 2015)

- The entry in (19) allows us to analyze (31a) correctly and concisely as follows:

(32) $[\![(31a)]\!] = \exists e.\ \text{unharmonious}(e) \wedge {}^*\text{agent}(e) = \bigoplus \text{organ.student}$

$\wedge\ e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\text{sustain}(e') \wedge \text{note}(\text{theme}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$

(There is an unharmonious event *e* whose agent is all the students and which consists entirely of note-sustaining events whose agents are individual students.)

- Building on insights by Schein (1993) and Kratzer (2000), we can also account for cumulative readings of *every* and *each*. Such configurations cause problems for the traditional analysis in (27).

(33) a. Three video games taught every quarterback two new plays.
                                                        (Schein 1993)
      b. Three copy editors caught every mistake (in the manuscript).
                                                        (Kratzer 2000)
      c. Two farmers sold each sheep to one customer.
                                                        (Thomas & Sudo 2016)

- My analysis of this reading is as follows:

(34) $[\![(33b)]\!] = \exists e. \, |{}^{*}\mathrm{agent}(e)| = 3 \wedge {}^{*}\mathrm{copy\text{-}editor}({}^{*}\mathrm{agent}(e)) \,\wedge$

$\quad {}^{*}\mathrm{theme}(e) = \bigoplus \mathrm{mistake} \wedge e \in {}^{*}\lambda e' \begin{pmatrix} {}^{*}\mathrm{catch}(e') \,\wedge \\ \mathrm{Atom}(\mathrm{theme}(e')) \end{pmatrix}$

(There is an event whose agents sum up to three copy editors, whose themes sum up to all the mistakes, and which consists of catching events with atomic themes.)

- It appears that *every* can never enter a cumulative relation with an argument in its syntactic scope (Champollion 2010a). For example, (35) does not have a cumulative reading (Kratzer 2000):

(35)   Every copy editor caught 500 mistakes.

- Bayer (1997) judges (36a) to be "clearly bizarre" because scripts cannot be written more than once, but reports that (36b) has a reading where every screenwriter in Hollywood contributed to the writing of the movie:

(36)   a.   Every screenwriter in Hollywood wrote *Gone with the Wind*.
        b.   *Gone with the Wind* was written by every screenwriter in Hollywood.

- Zweig (2008) reports that (37a) entails that each game was won by both teams at once, but (37b) has a cumulative reading:

(37)   a.   Every game was won by the Fijians and the Peruvians.
        b.   The Fijians and the Peruvians won every game.

- These facts are in line with what we would expect.

(38) $[\![(36a)]\!] = \exists e. \, {}^{*}\mathrm{agent}(e) = \bigoplus \mathrm{screenwriter} \,\wedge$

$\quad e \in {}^{*}\lambda e' \begin{pmatrix} {}^{*}\mathrm{write}(e') \wedge {}^{*}\mathrm{theme}(e') = [\![\text{Gone with the Wind}]\!] \,\wedge \\ \mathrm{Atom}(\mathrm{agent}(e')) \end{pmatrix}$

(39) $[\![(37a)]\!] = \exists e. \, {}^{*}\mathrm{theme}(e) = \bigoplus \mathrm{game} \,\wedge$

$\quad e \in {}^{*}\lambda e' \begin{pmatrix} {}^{*}\mathrm{win}(e') \wedge {}^{*}\mathrm{agent}(e') = [\![\text{the Fijians and the Peruvians}]\!] \\ \wedge \, \mathrm{Atom}(\mathrm{theme}(e')) \end{pmatrix}$

- By contrast, in (36b) and (37b), the syntactic scope of the *every*-phrase only includes the verb. For this reason, it does not distribute over the other argument, and a cumulative reading is possible.

(40) $\llbracket(36b)\rrbracket$
$= \exists e.\ ^*\text{theme}(e) = \llbracket\text{Gone with the Wind}\rrbracket\ \wedge$
$^*\text{agent}(e) = \oplus\text{screenwriter}\ \wedge\ e \in {}^*\lambda e' \begin{pmatrix} ^*\text{write}(e')\ \wedge \\ \text{Atom}(\text{agent}(e')) \end{pmatrix}$

(41) $\llbracket(37b)\rrbracket$
$= \exists e.\ ^*\text{agent}(e) = \llbracket\text{the Fijians and the Peruvians}\rrbracket\ \wedge$
$^*\text{theme}(e) = \oplus\text{game}\ \wedge\ e \in {}^*\lambda e' \begin{pmatrix} ^*\text{win}(e')\ \wedge \\ \text{Atom}(\text{theme}(e')) \end{pmatrix}$

- Kratzer 2000 claims that the availability of cumulative readings depends on the thematic role of the coargument of the *every*-phrase. According to her, cumulativity is only possible when the coargument plays the agent role. However, (36b), where the role of the coargument is theme, is a counterexample (Champollion 2010a).

Appendix

Answer to Exercise 1.1: Yes, on the assumption that cumulativity holds of *see* and of the theme relation, the verb phrase *see John* has cumulative reference. A seeing-John event is a seeing event whose theme is John. We therefore need to prove that the sum of any two seeing-John event is both a seeing event and an event whose theme is John. From cumulativity of *see*, we know that the sum of any two seeing events is a seeing event, so the sum of any two seeing-John events is a seeing event. From cumulativity of *theme*, the theme of the sum of any two events whose individual themes are John is the sum of their individual themes, then the theme of the sum of these events is the sum of John and John, which is John, given that the sum operation is idempotent.

Answer to Exercise 1.2: For *see two apples*, the proof does not go through because the theme of *see* is holistic and not incremental, that is, there is no meaning postulate like Incremental$_{theme}$(⟦see⟧). For *eat apples*, the proof does not go through because *apples* is not quantized (the sum of any two things in the denotation of *apples* is again in the denotation of *apples*).

Answer to Exercise 4.1: The star operator $^{*}\lambda e'$ is introduced through the D operator and takes scope over the predicate *dress* introduced by the theme. (12) does not directly require the theme of e to be a dress, though it requires e to consist of parts whose themes are dresses. This allows for the possibility that each girl wears a potentially different dress. The representation explicitly states that the dress-wearing events e' have pure atoms as agents, but not that these pure atoms are girls. However, this fact is entailed by cumulativity of thematic roles together with the assumption that the entities in the denotation of singular count nouns are atoms. By cumulativity of thematic roles, any entity x which is the agent of one of the dress-wearing events e' is a part of the agent of e. This agent is the sum of all girls. By definition of sum, x overlaps with a part of this agent. Being atomic, x can only overlap with y if it is a part of y. This means that

x is an atomic part of the girls. Given the background assumption that singular individuals like girls are mereological atoms, it follows that *x* is a girl. In this way, the distributive interpretation of (12) is correctly captured.

Answer to Exercise 4.2: $C = \{rodgers \oplus hammerstein, rodgers \oplus hart\}$

Answer to Exercise 4.3: We assume that the verbal predicate is closed under sum:

(26) $\forall e, e'[\text{write}(e) \wedge \text{write}(e') \rightarrow \text{write}(e \oplus e')]$

We also assume that the agent and theme relations are closed under sum:

(27)   a. $\forall e, e', x, x'[\text{agent}(e) = x \wedge \text{agent}(e') = x' \rightarrow \text{agent}(e \oplus e') = x \oplus x']$

        b. $\forall e, e', x, x'[\text{agent}(e) = x \ \wedge \ \text{theme}(e') = x' \rightarrow \text{theme}(e \oplus e') = x \oplus x']$

Bibliography

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 15(1). 5–16. https://doi.org/10.1007/bf00627432.

Balusu, Rahul. 2005. Distributive reduplication in Telugu. *36th Annual Meeting of the North East Linguistic Society (NELS 36)*. 39–53. http://semanticsarchive.net/Archive/mE0YTZkZ.

Balusu, Rahul & Karattuparambil A. Jayaseelan. 2013. Distributive quantification by reduplication in Dravidian. In *Strategies of quantification*, 60–86. Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199692439.003.0004.

Barker, Chris. 1992. Group terms in English: Representing groups as atoms. *Journal of Semantics* 9(1). 69–93. https://doi.org/10.1093/jos/9.1.69.

Bayer, Samuel L. 1997. *Confessions of a lapsed Neo-Davidsonian: Events and arguments in compositional semantics*. New York, NY: Garland. https://doi.org/10.4324/9780203055106.

Beavers, John. 2008. Multiple incremental themes and Figure/Path relations. *18th Semantics and Linguistic Theory Conference (SALT 18)*. 90–107. https://doi.org/10.3765/salt.v18i0.2495.

Beck, Sigrid & Arnim von Stechow. 2007. Pluractional adverbials. *Journal of Semantics* 24(3). 215–254. https://doi.org/10.1093/jos/ffm003.

Bennett, Michael R. & Barbara H. Partee. 1972. *Toward the logic of tense and aspect in English*. Indiana University Linguistics Club. https://doi.org/10.1002/9780470751305.ch4.

Blaheta, Don. 2003. *Binominal* each*: Evidence for a modified type system*. Providence, RI: Brown University MA thesis. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.7455&rep=rep1&type=pdf.

Brisson, Christine. 1998. *Distributivity, maximality and floating quantifiers*. New Brunswick, NJ: Rutgers University dissertation. http://ling.rutgers.edu/images/10.1.1.121.4641.pdf.

Brisson, Christine. 2003. Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26(2). 129–184. https://doi.org/10.1023/a:1022771705575.

Cable, Seth. 2014. Distributive numerals and distance distributivity in Tlingit (and beyond). *Language* 90(3). 562–606. https://doi.org/10.1353/lan.2014.0055.

Carlson, Gregory N. 1977. *Reference to kinds in English*. Amherst, MA: University of Massachusetts dissertation. http://semanticsarchive.net/Archive/jk3NzRlY.

Casati, Roberto & Achille C. Varzi. 1999. *Parts and places: The structures of spatial representation*. Cambridge, MA: MIT Press.

Champollion, Lucas. 2010a. Cumulative readings of *every* do not provide evidence for events and thematic roles. *Logic, Language and Meaning: 17th Amsterdam Colloquium*. 213–222. https://doi.org/10.1007/978-3-642-14287-1_22.

Champollion, Lucas. 2010b. *Parts of a whole: Distributivity as a bridge between aspect and measurement*. Philadelphia, PA: University of Pennsylvania dissertation. http://repository.upenn.edu/edissertations/958/.

Champollion, Lucas. 2011. Temporal prepositional phrases and implicit variables. Ms. University of Tübingen. http://www.champollion.com/temporal-pps-and-implicit-variables.pdf.

Champollion, Lucas. 2016a. Covert distributivity in algebraic event semantics. *Semantics and Pragmatics* 9(15). 1–65. https://doi.org/10.3765/sp.9.15.

Champollion, Lucas. 2016b. Overt distributivity in algebraic event semantics. *Semantics and Pragmatics* 9(16). 1–65. https://doi.org/10.3765/sp.9.16.

Champollion, Lucas. 2017. *Parts of a whole: distributivity as a bridge between aspect and measurement*. Vol. 66 (Oxford Studies in Theoretical Linguistics). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oso/9780198755128.001.0001.

Champollion, Lucas. 2019. Distributivity in formal semantics. *Annual Review of Linguistics* 5(1). 289–308. https://doi.org/10.1146/annurev-linguistics-011718-012528.

Champollion, Lucas. 2021. Distributivity, collectivity, and cumulativity. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.), *The Wiley Blackwell companion to semantics*, 1–38. Hoboken, NJ: Wiley. https://doi.org/10.1002/9781118788516.sem021.

Champollion, Lucas & Manfred Krifka. 2016. Mereology. In Maria Aloni & Paul Dekker (eds.), *Cambridge handbook of formal semantics* (Cambridge Handbooks in Language and Linguistics), chap. 13, 369–388. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9781139236157.014.

Cheng, Chung-Ying. 1973. Comments on Moravcsik's paper. In Jaakko Hintikka, Julius Moravcsik & Patrick Suppes (eds.), *Approaches to natural language: 1970 Stanford workshop on grammar and semantics*, 286–288. Dordrecht, Netherlands: Reidel. https://doi.org/10.1007/978-94-010-2506-5_14.

Choe, Jae-Woong. 1987. *Anti-quantifiers and a theory of distributivity*. Amherst, MA: University of Massachusetts dissertation. http://scholarworks.umass.edu/dissertations/AAI8727027/.

Choe, Jae-Woong. 1991. Georgian reduplication and a relational theory of distributivity. In Akira Ikeya (ed.), *Sixth Japanese-Korean joint conference on formal linguistics*. Tokyo, Japan: The Logico-Linguistic Society of Japan. https://dspace.wul.waseda.ac.jp/dspace/bitstream/2065/11793/1/JK6-4-25.pdf.

Deo, Ashwini Sharad & Maria Mercedes Piñango. 2011. Quantification and context in measure adverbs. *21st Semantics and Linguistic Theory Conference (SALT 21)*. 295–312. https://doi.org/10.3765/salt.v21i0.2614.

Dobrovie-Sorin, Carmen. 2014. Collective quantification and the homogeneity constraint. *24th Semantics and Linguistic Theory Conference (SALT 24)*. 453–472. https://doi.org/10.3765/salt.v24i0.2428.

van der Does, Jaap. 1993. Sums and quantifiers. *Linguistics and Philosophy* 16(5). 509–550. https://doi.org/10.1007/978-94-017-2706-8_6.

van der Does, Jaap & Henk J. Verkuyl. 1996. Quantification and predication. In Kees van Deemter & Stanley Peters (eds.), *Semantic ambiguity and under-specification*, 27–54. Stanford, CA: CSLI Publications. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.37.8829&rep=rep1&type=pdf.

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht, Netherlands: Reidel. https://doi.org/10.1007/978-94-009-9473-7.

Dowty, David R. 1987. Collective predicates, distributive predicates, and *all*. In Ann Miller & Zheng-sheng Zhang (eds.), *3rd Eastern States Conference on Linguistics (ESCOL)*, 97–115. Columbus, OH: Ohio State University.

Dowty, David R. & Belinda Brodie. 1984. The semantics of "floated" quantifiers in a transformationless grammar. *3rd West Coast Conference on Formal Linguistics (WCCFL 3)*. 75–90.

Erlewine, Michael Yoshitaka & Hadas Kotek. 2016. A streamlined approach to online linguistic surveys. *Natural Language and Linguistic Theory* 34(2). 481–495. https://doi.org/10.1007/s11049-015-9305-9. http://dx.doi.org/10.1007/s11049-015-9305-9.

Farkas, Donka F. 1997. Dependent indefinites. In Francis Corblin, Danièle Godard & Jean-Marie Marandin (eds.), *Empirical issues in formal syntax and semantics*, 243–268. Bern, Switzerland: Peter Lang. http://people.ucsc.edu/~farkas/papers/reduplication.pdf.

Ferreira, Marcelo. 2005. *Event quantification and plurality*. Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/33697.

Gawron, Jean Mark. 2005. Generalized paths. *15th Semantics and Linguistic Theory Conference (SALT 15)*. http://www-rohan.sdsu.edu/~gawron/salt_paper.pdf.

Gil, David. 1982. *Distributive numerals*. Los Angeles, CA: University of California dissertation.

Gil, David. 1989. Nouns, verbs, and quantification. Ms. Tel Aviv University.

Gillon, Brendan S. 1987. The readings of plural noun phrases in English. *Linguistics and Philosophy* 10(2). 199–219. https://doi.org/10.1007/bf00584318.

Gillon, Brendan S. 1990. Plural noun phrases and their readings: A reply to Lasersohn. *Linguistics and Philosophy* 13(4). 477–485. https://doi.org/10.1007/bf00630751.

Groenendijk, Jeroen, Theo Janssen & Martin Stokhof (eds.). 1984. *Truth, interpretation and information: selected papers from the third Amsterdam Colloquium*. Vol. 2 (Groningen-Amsterdam Studies in Semantics (GRASS)). Dordrecht, Netherlands: Foris. https://doi.org/10.1515/9783110867602.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford, UK: Blackwell Publishing.

Kratzer, Angelika. 2000. The event argument and the semantics of verbs, chapter 2. Ms. University of Massachusetts, Amherst. http://semanticsarchive.net/Archive/GU1NWM4Z/.

Kratzer, Angelika. 2001. The event argument and the semantics of verbs, chapter 4. Ms. University of Massachusetts, Amherst.

Kratzer, Angelika. 2008. On the plurality of verbs. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Event structures in linguistic form and interpretation*, 269–300. Berlin, Germany: de Gruyter. https://doi.org/10.1515/9783110925449.269.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.), *Semantics and contextual expression*, 75–115. Dordrecht, Netherlands: Foris. http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201989%20Nominal%20Reference.pdf.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters*, 29–53. Stanford, CA: CSLI Publications. http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201992%20Thematic%20Relations.PDF.

Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and grammar*, vol. 70 (Studies in linguistics and philosophy), 197–235. Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-011-3969-4_9.

Kroch, Anthony S. 1974. *The semantics of scope in English*. Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/13020.

Kuhn, Jeremy. 2014. Gather-*type predicates: Massiness over participants*. Presentation at the 45th Annual Meeting of the North East Linguistic Society (NELS 45). http://www.jeremykuhn.net/papers/Kuhn-gather-slides.pdf.

Kuhn, Jeremy. 2020. Gather/numerous as a mass/count opposition. *Natural Language Semantics* 28(3). 225–253. https://doi.org/10.1007/s11050-020-09163-x.

Landman, Fred. 1991. *Structures for semantics*. Kluwer.

Landman, Fred. 1996. Plurality. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 425–457. Oxford, UK: Blackwell Publishing. https://doi.org/10.1111/b.9780631207498.1997.00019.x.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures*. Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-011-4359-2.

Lasersohn, Peter. 1989. On the readings of plural noun phrases. *Linguistic Inquiry* 20(1). 130–134. http://www.jstor.org/stable/4178619.

Lasersohn, Peter. 1995. *Plurality, conjunction and events*. Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-015-8581-1.

Lasersohn, Peter. 1998. Generalized distributivity operators. *Linguistics and Philosophy* 21(1). 83–93. https://doi.org/10.1023/a:1005317815339.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Reiner Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use and interpretation of language*, 303–323. Reprinted in Link 1998a, Chapter 1. Also reprinted in , Portner & Partee 2002, Chapter 4. Berlin, Germany: de Gruyter. https://doi.org/10.1002/9780470758335.ch4.

Link, Godehard. 1987. Generalized quantifiers and plurals. In Peter Gärdenfors (ed.), *Generalized quantifiers*, 151–180. Reprinted in Link 1998a, Chapter 4. Berlin, Germany: Springer. https://doi.org/10.1007/978-94-009-3381-1_6.

Link, Godehard. 1997. Ten years of research on plurals—where do we stand? In Fritz Hamm & Erhard Hinrichs (eds.), *Plurality and quantification*, vol. 69 (Studies in Linguistics and Philosophy), 19–54. Reprinted in Link 1998a, Chapter 7. Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-017-2706-8_2.

Link, Godehard. 1998a. *Algebraic semantics in language and philosophy*. Stanford, CA: CSLI Publications.

Link, Godehard. 1998b. Je drei Äpfel—three apples each: Quantification and the German 'je'. In Godehard Link (ed.), *Algebraic semantics in language and philosophy*, 117–132. Stanford, CA: CSLI Publications.

Lønning, Jan Tore. 1987. Mass terms and quantification. *Linguistics and Philosophy* 10(1). 1–52. https://doi.org/10.1007/bf00603391.

Malamud, Sophia A. 2006a. (Non)maximality and distributivity: A decision theory approach. *16th Semantics and Linguistic Theory Conference (SALT 16)*. 120–137. https://doi.org/10.3765/salt.v16i0.2943.

Malamud, Sophia A. 2006b. *Semantics and pragmatics of arbitrariness*. Philadelphia, PA: University of Pennsylvania dissertation. http://people.brandeis.edu/~smalamud/Malamud-diss-single.pdf.

Malamud, Sophia A. 2012. The meaning of plural definites: A decision-theoretic approach. *Semantics and Pragmatics* 5(3). 1–58. https://doi.org/10.3765/sp.5.3.

Moltmann, Friederike. 1991. Measure adverbials. *Linguistics and Philosophy* 14(6). 629–660. https://doi.org/10.1007/bf00631962.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Jaakko Hintikka, Julius Moravcsik & Patrick Suppes (eds.), *Approaches to natural language: 1970 Stanford workshop on grammar and semantics*, 221–242. Dordrecht, Netherlands: Reidel. https://doi.org/10.1007/978-94-010-2506-5_10.

Müller, Ana & Esmeralda Negrão. 2012. On distributivity in Karitiana. In Patricia Cabredo Hofherr & Brenda Laca (eds.), *Verbal plurality and distributivity*, vol. 546 (Linguistische Arbeiten), 159–184. de Gruyter. https://doi.org/10.1515/9783110293500.159.

Partee, Barbara H., Alice G. B. ter Meulen & Robert Eugene Wall. 1990. *Mathematical methods in linguistics*. Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-009-2213-6.

Pontow, Carsten & Rainer Schubert. 2006. A mathematical analysis of theories of parthood. *Data & Knowledge Engineering* 59(1). 107–138. https://doi.org/10.1016/j.datak.2005.07.010.

Portner, Paul & Barbara H. Partee (eds.). 2002. *Formal semantics: the essential readings*. Blackwell. https://doi.org/10.1002/9780470758335.

Pratt, Ian & Nissim Francez. 2001. Temporal prepositions and temporal generalized quantifiers. *Linguistics and Philosophy* 24(2). 187–255. https://doi.org/10.1023/a:1005632801858.

Roberts, Craige. 1987. *Modal subordination, anaphora, and distributivity*. Amherst, MA: University of Massachusetts dissertation. http://scholarworks.umass.edu/dissertations/AAI8710499/.

Rooth, Mats. 1995. Indefinites, adverbs of quantification, and focus semantics. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, chap. 6, 265–299. Chicago, IL: University of Chicago Press.

Rothstein, Susan. 2004. *Structuring events*. Oxford, UK: Blackwell Publishing. https://doi.org/10.1002/9780470759127.

Rothstein, Susan. 2009. Individuating and measure readings of classifier constructions: Evidence from Modern Hebrew. *Brill's Annual of Afroasiatic Languages and Linguistics* 1(1). 106–145. https://doi.org/10.1163/187666309x12491131130783.

Safir, Ken & Tim Stowell. 1988. Binominal each. *18th Annual Meeting of the North East Linguistic Society (NELS 18)*. 426–450. http://www.linguistics.ucla.edu/people/stowell/Safir-Stowell_-_Binominal_Each_-_NELS_1988_ocr.pdf.

Scha, Remko. 1981. Distributive, collective and cumulative quantification. In Jeroen A.G. Groenendijk, Theo M.V. Janssen & Martin B.J. Stokhof (eds.), *Formal methods in the study of language*, vol. 2, 483–512. Reprinted in Groenendijk, Janssen & Stokhof 1984, page 131-158. Amsterdam, Netherlands: Mathematisch Centrum tract 136. https://doi.org/10.1515/9783110867602.131.

Schein, Barry. 1986. *Event logic and the interpretation of plurals*. Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/15107.

Schein, Barry. 1993. *Plurals and events*. Cambridge, MA: MIT Press.

Schwarzschild, Roger. 1991. *On the meaning of definite plural noun phrases*. Amherst, MA: University of Massachusetts dissertation.

Schwarzschild, Roger. 1996. *Pluralities*. Vol. 61 (Studies in Linguistics and Philosophy). Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-017-2704-4.

Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9(1). 67–110. https://doi.org/10.1111/j.1467-9612.2006.00083.x.

Simons, Peter. 1987. *Parts: A study in ontology*. Clarendon Press.

von Stechow, Arnim. 2002. Temporal prepositional phrases with quantifiers: Some additions to Pratt and Francez (2001). *Linguistics and Philosophy* 25(5-6). 755–800. https://doi.org/10.1023/a:1020872017811.

Szabolcsi, Anna. 2010. *Quantification*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511781681.

Tarski, Alfred. 1935. Zur Grundlegung der Booleschen Algebra. *Fundamenta Mathematicae* 24. Translated to English as Tarski 1956, 177–198. https://eudml.org/doc/212745.

Tarski, Alfred. 1956. On the foundations of the Boolean algebra. In Joseph Henry Woodger (ed.), *Logics, semantics, metamathematics. Papers from 1923 to 1938*, 320–341. Translation of Tarski 1935. Oxford, UK: Clarendon.

Taylor, Barry. 1977. Tense and continuity. *Linguistics and Philosophy* 1(2). 199–220. https://doi.org/10.1007/bf00351103.

Taylor, Barry. 1985. *Modes of occurrence*. Oxford, UK: Basil Blackwell.

Thomas, Guillaume. 2015. The present tense is not vacuous. *Journal of Semantics* 32(4). 685–747. https://doi.org/10.1093/jos/ffu010.

Thomas, Guillaume & Yasutada Sudo. 2016. *Cumulative readings of* each. Presentation at the Workshop on (Co-)Distributivity, Paris, February 11-12, 2016. https://www.ucl.ac.uk/~ucjtudo/pdf/eachParis2016.pdf.

Todorova, Marina, Karen Straub, William Badecker & Robert Frank. 2000. Aspectual coercion and the online computation of sentential aspect. In Lila R. Gleitman & Aravind K. Joshi (eds.), *22nd Annual Conference of the Cognitive Science Society*, 3–8. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.23.8886&rep=rep1&type=pdf.

Varzi, Achille C. 2010. Mereology. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Stanford, CA: Metaphysics Research Lab, Stanford University. http://plato.stanford.edu/archives/spr2010/entries/mereology/.

Verkuyl, Henk J. 1972. *On the compositional nature of the aspects*. Dordrecht, Netherlands: Reidel. https://doi.org/10.1007/978-94-017-2478-4.

Verkuyl, Henk J. & Jaap van der Does. 1991. *The semantics of plural noun phrases*. Vol. LP-91-07 (ITLI prepublication series for logic, semantics and philosophy of language). Amsterdam, Netherlands: University of Amsterdam. https://www.illc.uva.nl/Research/Publications/Reports/LP-1991-07.text.pdf.

Winter, Yoad. 2001. *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge, MA: MIT Press.

Zimmermann, Malte. 2002. *Boys buying two sausages each: On the syntax and semantics of distance-distributivity*. Amsterdam, Netherlands: University of Amsterdam dissertation. http://hdl.handle.net/11245/1.194584.

Zimmermann, Malte. 2008. Quantification in Hausa. In Lisa Matthewson (ed.), *Quantification: Universals and variation*, vol. 64 (North Holland Linguistic Series: Linguistic Variations), chap. 11, 415–475. Bingley, UK: Emerald.

Zucchi, Sandro & Michael White. 2001. Twigs, sequences and the temporal constitution of predicates. *Linguistics and Philosophy* 24(2). 223–270. https://doi.org/10.1023/a:1005690022190.

Zweig, Eytan. 2008. *Dependent plurals and plural meaning*. New York, NY: New York University dissertation. http://semanticsarchive.net/Archive/GJhMzhhZ/.

Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32(4). 353–407. https://doi.org/10.1007/s10988-009-9064-3.