The common core of distributivity, aspect and measurement*

Lucas Champollion
New York University

champollion@nyu.edu

CNRS, Paris 7 — January 16^{th}, 2014

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*

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).)

*This research is part of a dissertation written at Penn, Stanford, and PARC, and was further developed during a postdoc in Tübingen. I am grateful to all these institutions. Special thanks to my advisor, Cleo Condoravdi; my committee members, Aravind Joshi, Maribel Romero, Florian Schwarz; to Sigrid Beck, Danny Bobrow, Beth Levin, Lauri Karttunen, Paul Kiparsky, Tony Kroch, Chris Potts, Annie Zaenen, and many others.

Examples:

- Volume is monotonic $\rightsquigarrow$ *thirty liters of water*
- Temperature is not monotonic $\rightsquigarrow$ **thirty degrees Celsius of water*
- What about height? It had better be monotonic: $\rightsquigarrow$ *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**.

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.

2 Answer strategy

For-adverbials are most commonly associated with the telic/atelic opposition.

- Atelic predicates: walk, sleep, eat apples, run, run towards the store
  ($\approx$ *as soon as you start X-ing, you have already X-ed*)
- Telic predicates: build a house, eat ten apples, run to the store
  ($\approx$ *you need to reach a set terminal point in order to have X-ed*)

(5)	a.	John **ran** for five minutes.	*atelic*
	b.	*John **ran to the store** for five minutes.	*telic*
(6)	a.	John **ate apples** for an hour.	*atelic*
	b.	*John **ate ten apples** for an hour.	*telic*

Plan of the talk:

- Introduce *stratified reference*, which generalizes the telic/atelic contrast.
- Derive the restriction on measure functions from this concept.
- Apply this concept to two puzzles in distributivity.

3 The aspect puzzle

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 and Partee, 1972; Dowty, 1979). Here's an event-based version of it:

(7) SUBINT(P) $=_{def}$
$\forall e[P(e) \rightarrow \forall i[i < \text{runtime}(e) \rightarrow \exists e'[P(e') \wedge e' < e \wedge i = \text{runtime}(e')]]]$
(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.)

On a Dowty-style account, *for*-adverbials presuppose the subinterval property.

3.1 Problems with the subinterval property

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

(8) 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:

(9) The Chinese people have created abundant folk arts ... passed on from generation to generation for thousands of years.[1]

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

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

(11) The police blocked streets for miles around [the museum].[2]

3.2 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.

[1] Attested example (`http://www.twinbridge.com/detail.aspx?ID=315`). Nov 15, 2010.
[2] Attested example, New York Times, February 24, 2009. Thanks to Cleo Condoravdi.

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

(12) $\forall e[\text{waltz}(e) \rightarrow \forall i[i < \text{runtime}(e) \rightarrow \exists e'[\text{waltz}(e') \wedge e' < e \wedge i = \text{runtime}(e')]]]$
(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.)

Let ε be a function that tells us what counts as very small: $\varepsilon(\lambda t[\text{hours}(t) = 1])(t')$ is true just in case t' is very small with respect to one hour.

We want to be able to say:

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

To express this formally, we use the star operator from Link (1987).

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

With the star operator, we can express (13) as follows:

(14) $$\forall e[\text{waltz}(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{l} \text{waltz}(e') \wedge \\ \varepsilon(\lambda t[\text{hours}(t) = 1])(\text{runtime}(e')) \end{array} \right)]$$

Let us say that *waltz* has **stratified reference** (SR) with respect to the dimension *runtime* and the granularity $\varepsilon(\lambda t[\text{hours}(t) = 1])$ ("very short time interval") just in case (14) above is true.

(15) **Stratified reference (Example)**
Let "$\text{SR}_{runtime,\ \varepsilon(\lambda t[hours(t) = 1])}(\lambda e[\text{waltz}(e)])$" abbreviate (14).

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

(16) **Stratified reference (Definition)**
$$\text{SR}_{f,\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow x \in {}^*\lambda y \left(\begin{array}{l} P(y) \wedge \\ \varepsilon(K)(f(y)) \end{array} \right)]$$

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:

(17) waltz for an hour
Satisfied presupposition:
$$\forall e[\text{waltz}(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{l} \text{waltz}(e') \wedge \\ \varepsilon(\lambda t[\text{hours}(t) = 1])(\text{runtime}(e')) \end{array} \right)]$$

(Every waltzing event consists of waltzing subevents whose runtimes are very small compared to an hour.)

Note: I assume following Zweig (2008) and others that ⟦eat apples⟧ = "eat one or more apples" (the "two or more" meaning component is an implicature)

(18) eat apples for three hours
Satisfied presupposition:
$\forall e[[\![\text{eat apples}]\!](e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{l} [\![\text{eat apples}]\!](e') \wedge \\ \varepsilon(\lambda t[\text{hours}(t) = 3])(\text{runtime}(e')) \end{array} \right)]$
(Every event in which one or more apples are eaten consists of subevents in which one or more apples are eaten and whose runtimes are very small compared to three hours.)

(19) *eat ten apples for three hours
Failing presupposition:
$\forall e[[\![\text{eat ten apples}]\!](e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{l} [\![\text{eat ten apples}]\!](e') \wedge \\ \varepsilon(\lambda t[\text{hours}(t) = 3])(\text{runtime}(e')) \end{array} \right)]$
(Every eating-ten-apples event consists of eating-ten-apples subevents whose runtimes are very small compared to three hours.)

4 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:

(20) 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.

(21) *drive for thirty miles per hour
Failing presupposition: $\text{SR}_{\text{speed},\varepsilon([\![\text{thirty mph}]\!])}([\![\text{drive}]\!])$
(Every driving event consists of driving subevents whose speeds are very small compared to thirty mph.)

Now we transfer this idea to pseudopartitives.

4.1 Baseline examples

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

(22) run for three hours / three hours of running
Satisfied presupposition: $\mathrm{SR}_{\mathrm{runtime},\varepsilon(\llbracket\text{three hours}\rrbracket)}(\llbracket\text{run}\rrbracket)$
(Every running event consists of running subevents whose runtimes are very small compared to three hours.)

The dimension parameter is the appropriate measure function.

(23) thirty liters of water
Satisfied presupposition: $\mathrm{SR}_{\mathrm{volume},\varepsilon(\llbracket\text{thirty liters}\rrbracket)}(\llbracket\text{water}\rrbracket)$
(Every water amount consists of water parts whose volumes are very small compared to thirty liters.)

4.2 Temperature in pseudopartitives

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

(24) *thirty degrees Celsius of water
Failing presupposition: $\mathrm{SR}_{\mathrm{temperature},\varepsilon(\llbracket\text{thirty degrees Celsius}\rrbracket)}(\llbracket\text{water}\rrbracket)$
(Every water amount consists of water parts whose temperatures are very low compared to thirty degrees Celsius.)

4.3 The problematic snow example

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

(25) five feet of snow
Satisfied presupposition: $\mathrm{SR}_{\mathrm{height},\varepsilon(\llbracket\text{five feet}\rrbracket)}(\llbracket\text{snow}\rrbracket)$
(Every snow amount consists of snow parts whose heights are very small compared to five feet.)

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.

5 Distributivity in English

Stratified reference also clarifies the *distributivity-collectivity* opposition.

5.1 Zweig's puzzle of cumulative readings

Zweig (2008) notes that *all* cannot give rise to cumulative (scopeless) readings:

(26) a. Three safari participants saw thirty zebras.
Available 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 reading: Each safari participant saw at least one zebra, and thirty zebras were seen overall.

Exception: dependent-plural readings, which can be seen as cumulative:

(27) a. Three safari participants saw zebras.
Available 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 *reading:* Each safari participant saw at least one zebra, and at least two zebras were seen overall.

Zweig's Puzzle: What is the relevant semantic distinction between *see zebras* and *see thirty zebras*?

Aspect Puzzle (reminder): What is the relevant semantic distinction between *eat apples* (atelic) and *eat thirty apples* (telic)?

5.2 Stratified reference as distributivity

All requires distributivity (Dowty, 1987):

(28) a. All the children smiled. ⇒ Each child smiled.
b. *All the children are numerous.
c. All the committees are numerous. ⇒ Each committee is numerous.

Assumption: *All* imposes a constraint which is analogous to the presupposition of *for*-adverbials, but the parameters are instantiated differently.

(29) **Presupp. of *for 1h*:** $\forall e[\mathrm{VP}(e) \rightarrow e \in {}^*\lambda e' \begin{pmatrix} \mathrm{VP}(e') \wedge \\ \varepsilon(\lambda t[\mathrm{hours}(t) = 1])(\mathrm{runtime}(e')) \end{pmatrix}]$
(Every VPing event consists of one or more VPing events whose *runtimes* are *very short* compared to an hour.)

(30) **Presupposition of *all*:** $\forall e[\text{VP}(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{l} \text{VP}(e') \wedge \\ \text{Atom}(\text{ag}(e')) \end{array} \right)]$
(Every VPing event consists of one or more VPing events whose *agents* are *atoms*.)

Baseline example:

(31) All the children smiled.
Presupposition: $\forall e[\text{smile}(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{l} \text{smile}(e') \wedge \\ \text{Atom}(\text{ag}(e')) \end{array} \right)]$
(Every smiling event consists of one or more smiling events whose agents are atoms. This entails that each child smiled.)

5.3 Zweig's puzzle

(32) 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 (32a) as a presupposition failure:

(33) **Failing presupposition:** $\text{SR}_{\text{agent, Atom}}([\![\text{see thirty zebras}]\!])$
(Every see-thirty-zebras event consists of subevents with atomic agents and in each of which thirty zebras are seen.)

The cumulative reading of (32b) is available, though:

(34) **Satisfied presupposition:** $\text{SR}_{\text{agent, Atom}}([\![\text{see zebras}]\!])$
(Every event in which at least one zebra is seen consists of subevents with atomic agents and in each of which at least one zebra is seen.)

Answers.
What is the relevant distinction between *eat apples* and *eat ten apples*?

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

What is the relevant distinction between *see zebras* and *see thirty zebras*?

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

6 Distributivity across languages

- Distributivity is often expressed with distance-distributive (DD) items, e.g. adnominal *each*:

(35) The children have seen two monkeys each.

- In some languages[3], DD items have a wider range of meanings than in others[4] (Zimmermann, 2002):
  - English-type DD items can only distribute over individuals.
  - German-type DD items can also distribute over occasions (= salient chunks of time or space).

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

(37) Hans hat jeweils zwei Affen gesehen. *German*
Hans has EACH two monkeys seen.
'Hans has seen two monkeys each time'

(38) 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 each time'

Zimmermann's Generalization: If a DD item can also be used as a distributive determiner, it lacks the occasion reading.

Questions:
- How can we capture the semantic variation among DD items?
- What is the reason for Zimmermann's generalization?

6.1 Capturing the semantic variation

- We have seen that distributivity is a parametrized property.
  - The **dimension** parameter indicates the domain of distributivity: e.g. a thematic role in the case of determiner *each*, or time/space in the case of *for*-adverbials (*for an hour*, *for a mile*).

[3]e.g. German, Czech, Bulgarian, Japanese, Korean
[4]e.g. English, French, Dutch, Norwegian, Icelandic, Italian, Russian, Turkish

  - The **granularity** parameter indicates the size of the entities over which we distribute: e.g. atoms or amounts of space or time.

- 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.

- *Each* and *jeweils* include two versions of the distributivity operator.

  - *Each* includes an atomic distributivity operator (Link, 1987), which can only distribute over count domains (granularity = atomic).
  - *Jeweils* includes a cover-based distributivity operator (Schwarzschild, 1996), which can also distribute over noncount domains (granularity = contextual).

- Here is what contextually-valued distributivity looks like:

(39) The shoes cost $50. *(per pair)* (Lasersohn, 1998)

- The limitation to supporting context is also present in *jeweils*, as we have seen above.

6.2 Explaining the semantic variation

- Recall **Zimmermann's Generalization**: If a DD item can also be used as a distributive determiner, it lacks the occasion reading.

- The way Zimmermann explains his generalization can be reconceptualized in terms of stratified reference.

- In English etc., adnominal and determiner *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.

- In German etc., *jeweils* cannot be used as a determiner, so it is free to have different properties from determiners.

Answers.

- How can we formally capture the semantic variation among DD items?
  - The variation emerges from "atomic" vs. "contextual" settings of the granularity parameter.
- What is the reason for Zimmermann's generalization?
  - DD items which also function as determiners carry their "atomic" setting over from them.

7 Summary

We have used a parametrized higher-order property, stratified reference, to

- improve on a notion developed for aspect
- solve a problem in the study of measurement
- and to get a new perspective on distributivity.

Emerging bigger picture: (Champollion, 2010, 2012)

(40) *atelic : telic :: mass/plural : count :: distributive : collective*

Distributivity as a property with two parameters:

- dimension: runtime, spatial extent, measure functions, thematic roles ...
- granularity: atomic (*all*, *each*), contextual (*for*, *jeweils*, pseudopartitives)

References

Bennett, M. R. and Partee, B. H. (1972). *Toward the logic of tense and aspect in English*. Indiana University Linguistics Club.

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

Champollion, L. (2012). A cover-based account of distance distributivity. In Aloni, M., Roelofsen, F., Sassoon, G. W., Schulz, K., Kimmelman, V., and Westera, M., editors, *Proceedings of the 18th Amsterdam Colloquium*, pages 251–260. To be presented at 2011 Amsterdam Colloquium.

Dowty, D. R. (1979). *Word meaning and Montague grammar*. Reidel, Dordrecht, Netherlands.

Dowty, D. R. (1987). Collective predicates, distributive predicates, and *All*. In *Proceedings of the Third Eastern States Conference on Linguistics (ESCOL 3)*, pages 97–115, Columbus, OH. The Ohio State University.

Gawron, J. M. (2005). Generalized paths. In *Proceedings of the Fifteenth Conference on Semantics and Linguistic Theory (SALT XV)*.

Krifka, M. (1998). The origins of telicity. In Rothstein, S., editor, *Events and grammar*, pages 197–235. Kluwer, Dordrecht, Netherlands.

Lasersohn, P. (1998). Generalized distributivity operators. *Linguistics and Philosophy*, 21(1):83–93.

Link, G. (1987). Generalized quantifiers and plurals. In Gärdenfors, P., editor, *Generalized Quantifiers: Linguistic and Logical Approaches*, pages 151–180. Reidel, Dordrecht, Netherlands. Reprinted in Link (1998), Chapter 4.

Link, G. (1998). *Algebraic semantics in language and philosophy*. CSLI Publications, Stanford, CA.

Schwarzschild, R. (1996). *Pluralities*. Kluwer, Dordrecht, Netherlands.

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

Taylor, B. (1977). Tense and continuity. *Linguistics and Philosophy*, 1:199–220.

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

Zweig, E. (2008). *Dependent plurals and plural meaning*. PhD thesis, New York University, New York, NY.