

# Eliminating EARLIEST: a general semantics for *before* and *after*<sup>1</sup>

Jessica RETT — *University of California, Los Angeles*

**Abstract.** Despite recent proposals, there are still empirical gaps in the study of the semantics of temporal relations, including the precise truth conditions of *before* and *after* constructions. Theoretically, there is disagreement about whether the constructions involve aspectual coercion operators, like EARLIEST, that privilege certain bounds over others (Beaver and Condoravdi, 2003; Condoravdi, 2010); or whether non-veridical interpretations of *before* constructions are the result of intensionality or semantic underspecification (Krifka, 2010b). In this paper, I focus on the truth conditions of *before* and *after* constructions as they are conditioned by different aspectual classes across languages, and propose a reformulation of the semantics of *before* and *after* that characterizes them as antonymic in a particular way inspired by the treatment of comparatives in the degree-semantic literature. I argue that the result is a more empirically comprehensive and explanatory theory of relations between ordered plurals in general.

**Keywords:** temporal relations, aspect, maximality, interval semantics, degree semantics.

## 1. Introduction

It has been clear since Elizabeth Anscombe's (1964) seminal paper that *before* and *after* constructions differ from one another in asymmetric ways. Her observations about the differences between the two regarded their (anti-)transitivity, (anti-)symmetry, and potential for 'alternative verifications'; in subsequent discussion, linguists tend to cast these same differences in terms of NPI-licensing, veridicality, and ambiguity (Heinämäki, 1974; Ogihara, 1995). And while Anscombe concluded that these differences indicated *before* and *after* are not lexical converses, the subsequent consensus has been to treat them as lexical converses and to explain the asymmetric differences in *before* and *after* constructions by appealing to sentence-level phenomena, like aspectual coercion (Heinämäki, 1974); metaphysical differences in precedence/subsequence relations (Beaver and Condoravdi, 2003; Condoravdi, 2010); or broad pragmatic effects (Krifka, 2010b).

A series of recent semantic treatments of *before* and *after* have focused principally on accounting for intuitions that the internal arguments of *after* (but not necessarily for *before*) constructions are entailed, and that NPIs are licensed in the internal arguments of all *before* constructions (Beaver and Condoravdi, 2003; Condoravdi, 2010; Krifka, 2010b), but only a restricted subset of *after* constructions. I argue that these analyses capture these veridicality and NPI facts at the expense of properly characterizing their truth conditions.

In what follows, I present a more in-depth discussion of the truth conditions imposed by *be-*

---

<sup>1</sup>This paper has benefited immensely from participants in my 2019 graduate seminar, as well as discussion with Daniel Altshuler, Dylan Bumford, Cleo Condoravdi, Sam Cumming, Gabe Greenberg, Jeff King, Nathan Klinedinst, Maziar Toorsarvandani, Yael Sharvit, 2017 audiences at the UConn Logic Colloquium and MIT LingLunch, and of course the 2019 SuB audience. I am also indebted to Nicoletta Loccioni for her help as a research assistant, and to several native speaker consultants: Maayan Abenina-Adar; Nikos Angelopoulos; Daniela Čulinović; Philippe Cote-Boucher; Moshe Elyashiv Bar-Lev; David Erschler; Alexander Goebel; Zsofia Gyarmathy; Marju Kaps; Petr Kusliy; Isabel Lin; Nicoletta Loccioni; Daniel Margulis; Victoria Mateu; Yoko Mizuta; Maša Močnik; Sozan Okzan; Una Stojnić; Yu Tanaka; and Jos Tellings. Funding was provided by a grant from the UCLA Faculty Committee on Research.

*fore* and *after*, as well as a cross-linguistic typological survey. My goal is to provide information about which asymmetrical differences between *before* and *after* are language-specific (and putatively semantic) and which are language-general (and therefore more appropriately characterized as attributable to pragmatics or metaphysics). I then propose an account that characterizes *before* and *after* as antonyms, and couples independently motivated aspectual coercion operations with a domain-general notion of maximal informativity. I argue that this account correctly characterizes the truth conditions of *before* and *after* sentences, and can be implemented, as other theories have, to account for NPI and veridicality issues.

## 2. Asymmetries between *before* and *after*

While Anscombe (1964) discussed several asymmetries between *before* and *after* constructions, the more recent consensus (starting with Heinämäki, 1974; Ogihara, 1995) is that the temporal relations differ in two main respects: *after* – but generally not *before* – entails its internal argument (the ‘veridicality asymmetry’); and *before* – but generally not *after* – licenses NPIs in its internal argument (the ‘NPI asymmetry’). I’ll present these central differences before discussing the truth conditions of *before* and *after* constructions in Section 2.3.

### 2.1. The veridicality asymmetry

The canonical illustration of the veridicality asymmetry between *before* and *after* is in (1) (from Beaver and Condoravdi, 2003).

- (1) a. Mozart died before he finished the Requiem.  $\nrightarrow$  Mozart finished the Requiem.  
 b. Mozart died after he finished the Requiem.  $\rightarrow$  Mozart finished the Requiem.

In (1) the proposition associated with internal argument – the clause ‘he finished the Requiem’ – is entailed by (1b) but not by (1a). In event-semantic terms, the eventuality associated with the internal argument is actualized in (1b) but not in (1a). Several, beginning with Heinämäki (1974), differentiate between two non-veridical interpretations: those compatible with a non-actualized embedded event (a ‘neutral’ reading), and those that entail a non-actualized embedded event (a ‘counterfactual’ reading).

There is a surprising amount of variation of judgment regarding which *before* sentences are veridical and which are not. Heinämäki (1974) argues that veridicality is conditioned by real-world knowledge triggered by the main clause verb (e.g. *died*); by the aspect of the main clause (claiming non-veridical constructions must have non-stative main clause verbs); by NPIs (compare *I left the country before something/anything happened*); and by ellipsis in the embedded clause (compare *John left before Bill did* with *John left before Bill*). Many of these claims have been challenged; Condoravdi (2010) goes so far as to argue that the (non-)veridicality of a *before* sentence is so context-sensitive that it cannot even be reliably tested for its status as an implicature or a presupposition.

### 2.2. The NPI asymmetry

There is, generally speaking, an asymmetry between *before* and *after* with respect to their ability to license negative polarity items (NPIs) in their embedded argument (Heinämäki, 1974; Ogihara, 1995). This is exemplified in (2).



- b. culminated ME *after* culminated EE:  
John met Mary after she climbed the mountain.  $\succ$  final
- (5) **ambiguous *before* and *after* sentences**
- a. culminated ME *before* culminated EE:  
John met Mary before she climbed the mountain.  $\prec$  initial,  $\prec$  final
- b. culminated ME *after* process EE:  
John met Mary after she was president.  $\succ$  initial,  $\succ$  final

The sentences in (4) are clearly unambiguous: in order for (4a) to be true, John needs to have met Mary before her first day in office; the sentence is false if he meets her halfway through her term. And in order for (4b) to be true, John needs to have met Mary after she summited the mountain; the sentence is false if he meets her halfway up the mountain, i.e. after she has begun climbing the mountain. While I've illustrated these truth conditions using a stative EE (for (4a)) and a culminated process EE (for (4b)), the intuitions should extend to process or activity EEs and culminations or accomplishment EEs, respectively.

In contrast the sentences in (5) are ambiguous. This has been previously acknowledged Heinämäki (1974); Condoravdi (2010) for (5b), but as far as I can tell, the ambiguity in (5a) has gone unnoticed. It nevertheless exists: in English, at least, (5a) is compatible with a scenario in which John and Mary meet halfway up the mountain, i.e. before she has summited (the  $\prec$  final interpretation), and it is also compatible with a scenario in which John and Mary meet on the bus on the way to the base of the mountain (the  $\prec$  initial interpretation). Similarly, (5b) is compatible with a scenario in which John met Mary after she'd retired from politics (the  $\succ$  final interpretation), as well as a scenario in which the meeting occurred while she was in office (the  $\succ$  initial interpretation).

It is worthwhile acknowledging the entailment relation between each of these readings. For the *before* sentence in (5a),  $\prec$  initial entails  $\prec$  final; for the *after* sentence in (5b),  $\succ$  final entails  $\succ$  initial. One might nevertheless be tempted to characterize the available readings in (5) as a matter of semantic underspecification rather than ambiguity. (Anscombe (1964) explicitly makes this argument, albeit with diagnostics that don't seem compatible with how linguists today characterize the difference.) But I characterize it as an ambiguity for several reasons: first, native speakers report the intuition that the two readings are distinct, i.e. available in different contexts; and second, as I will mention in Section 2.4, languages use aspectual marking to disambiguate between the two interpretations in a way that mirrors aspectual coercion, which is how Heinämäki (1974) and Condoravdi (2010) viewed the ambiguity in (5b).

In sum, given (4) and (5), we see the truth conditions illustrated in Table 1.<sup>2</sup>

Notably, Table 1 shows no clear bias across the spectrum of *before* and *after* constructions towards the initial or earliest point of the embedded eventuality. In other words, Table 1 offers no justification for an EARLIEST operator in the absence of a LATEST operator (cf. Beaver and Condoravdi, 2003), or an *earliest* operator coupled with a MAX operator that predicts an additional *after* ambiguity but not an additional *before* ambiguity (cf. Condoravdi, 2010).

<sup>2</sup>Although see Section 4 for a discussion of some subtleties in the truth conditions of constructions in which both the ME and EE are states or non-culminating processes.

	process EE	example	culmination EE	example
<i>A before B</i>	$\prec$ initial	(4a)	$\prec$ initial $\prec$ final	(5a)
<i>A after B</i>	$\succ$ initial $\succ$ final	(5b)	$\succ$ final	(4b)

Table 1: truth conditions of *before* and *after* sentences

I will briefly discuss cross-linguistic variation in the semantics of these constructions before evaluating these theories in a little more detail in Section 2.5.

#### 2.4. Cross-linguistic variation in *before* and *after* constructions

I conducted an informal typological survey on the syntax and semantics of *before* and *after* constructions. It surveyed 17 languages (including English) from seven different language families, listed below.

- (6) **Afroasiatic:** Hebrew; **Austronesian:** Tagalog; **Indo-European:** Dutch, English, German (*Germanic*); Greek (*Hellenic*); French, Italian, Spanish (*Romance*); Russian, Serbo-Croatian, Slovenian (*Slavic*); **Japonic:** Japanese; **Sino-Tibetan:** Mandarin; **Turkic:** Turkish; **Uralic:** Estonian, Hungarian

The survey strongly suggests that the veridicality and NPI asymmetries are universal. Speakers of all languages reported a veridicality asymmetry for sentences like the ones in (1). And each language either didn't license NPIs under temporal prepositions (e.g. Greek, Spanish) or only licensed NPIs under *before* (not *after*).<sup>3</sup> This, in turn, suggests that both asymmetries should receive a treatment that does not predict cross-linguistic variation.

The universality of the truth conditions of *before* and *after* constructions is a little harder to determine, in part for several reasons that are rooted firmly in morphosyntax: i) as demonstrated in del Prete (2008), languages can in principle use degree-based 'earlier' and 'later' strategies for expressing temporal precedence and succession, rather than the temporal relation strategies encoded in *before* and *after*; ii) as Sharvit (2014) shows, languages can differ in how they interpret past tense in embedded clauses; and iii) perhaps surprisingly, there is tentative evidence that languages also differ in whether they allow stative EEs in *before* and *after* constructions.

Regarding the latter: while English licenses stative EEs under both *before* and *after*, but other languages (like Russian and Japanese) were reported to disallow stative EEs under either (7).<sup>4</sup>

- (7) a. ??Anja vstretila Dinu do togo, kak ona byla svobodna.  
Anna met Dina until then how she was free  
Intended: 'Anna met Dina before she was single.'
- b. ??Anja vstretila Dinu posle togo, kak ona byla svobodna.  
Anna met Dina after then how she was free  
Intended: 'Anna met Dina after she was single.'

*Russian*

And still other languages (like Dutch and Hungarian) appear to disallow stative EEs, but only

<sup>3</sup>Although see del Prete (2008) for an interesting discussion of some complications in Italian.

<sup>4</sup>My Russian consultants commented that each would be acceptable with an embedded inchoative marker.

under *after* (not *before*). For these languages, consultants reported that the *after* sentences are degraded without an inchoative marker, and were unable to offer a characterization of its truth conditions without an inchoative marker. In contrast, I found no language that disallowed stative EEs under *before* but not *after*.

- (8) a. Mary azelőtt találkozott Johnnal, hogy egyetemista volt.  
 Mary before met John.with that university.student was  
 ‘Mary met John before he was a university student.’  
 b. ??Mary azut’an találkozott Johnnal, hogy egyetemista volt.  
 Mary after met John.with that university.student was  
 Intended: ‘Mary met John after he was a university student.’ *Hungarian*

For languages without these morphosyntactic differences, I found some evidence of truth-conditional variation. The ambiguity reported in English for *before* sentences with culmination EEs (e.g. (5a)) does not seem to be universal; Dutch, Hungarian, Italian and Tagalog consultants reported that such sentences unambiguously received an  $\prec$  initial interpretation.

- (9) Nakilala ni-Mary si-John bago siya um-akyat sa bundok.  
 PFV.TV-meet GEN-Mary SUBJ-John before SUBJ.3SG PFV.AV-climb OBL mountain  
 ‘Mary met John before he climbed to the top of the mountain.’ *Tagalog*,  $\prec$  initial

And the ambiguity reported in English for *after* sentences with process EEs (e.g. (5b)) does not seem universal, either; in German and Turkish, these sentences were reported to be unambiguously  $\succ$  final.

- (10) Mary hat John getroffen, nachdem er Single war.  
 Mary had-3SG John met after he single was-3SG  
 ‘Mary met John after he was single.’ *German*,  $\succ$  final

Finally, languages that overtly mark aspectual alternations in EE clauses were found to morphologically differentiate between the two readings reported for the constructions in (5). For instance, Serbo-Croatian allows for a perfective/imperfective alternation in embedded temporal clauses, and this alternation (when accepted) seems to condition the interpretation. The perfective version in (11b) has a  $\prec$  final interpretation, and that the imperfective version – to the extent it’s acceptable without an overt inchoative marker – receives a  $\prec$  initial interpretation.

- (11) a. Mary je srela Johna pre nego ?to se  
 Mary-NOM is-PRES-3FS met-PP-3FS John-ACC before than PTCL REFL  
 peo na vrh planine.  
 climb-IMP-3MS on top-ACC mountain-GEN.  
 ‘Mary met John before he climbed to the top of the mountain.’  $\prec$  initial  
 b. Mary je srela Johna pre nego ?to se  
 Mary-NOM is-PRES-3FS met-PP-3FS John-ACC before than PTCL REFL  
 popeo na vrh planine.  
 climb-PP-3MS on top-ACC mountain-GEN.  
 ‘Mary met John before he climbed to the top of the mountain.’  $\prec$  final

In Tagalog, this difference is conditioned by the neutral, non-culminating perfective ( $\prec$  initial) and the culminating perfective (ability-and-involuntary-action, AIA,  $\prec$  final; Dell (1983)).



Unfortunately, while the combination of these three operators correctly predicts most of the possible interpretations of *before* and *after* constructions illustrated in Table 1, it cannot predict one: namely the ‘ $\prec$  initial’ interpretation of *before* sentences with culmination EEs, like *John met Mary before she climbed the mountain*, as in (5a). The embedded eventuality, by virtue of having a telos, is coerced into a single point corresponding to its telos, after which the difference between EARLIEST and MAX is trivial. These predictions are illustrated in Table 2.

	process EE	B&C account	culmination EE	B&C account
<i>A before B</i>	$\prec$ initial	EARLIEST or MAX	$\prec$ initial $\prec$ final	<b>not possible</b> ITOP
<i>A after B</i>	$\succ$ initial $\succ$ final	EARLIEST MAX	$\succ$ final	ITOP

Table 2: B&C predictions for *before* and *after* truth conditions

It is possible in principle for B&C to add a fourth operator, ITOP2, to take a culminated eventuality from its runtime to its initial point, but this innovation would incorrectly predict that *after* sentences with culminated EEs are similarly ambiguous, which they aren’t. And, quite subjectively, the theory seems complicated enough with only three operators in its inventory.

The B&C account of NPI-licensing and veridicality is particularly well done in Condoravdi (2010), who includes an extensive discussion of veridicality (as e.g. a semantic or pragmatic presupposition) and NPI-licensing. She concludes that the NPI asymmetry is predicted given the notion of Strawson-entailment as a licensing condition (von Stechow, 1999), and can and should be disentangled from the (non-)veridicality of a *before* construction. Specifically, in the account, NPIs are licensed in *before* constructions because EARLIEST creates a Strawson-downward-entailing environment in the embedded context, and this remains true on an intensional level regardless of whether the EE occurs in the actual world.

The optionality between EARLIEST and MAX for *after* constructions has the additional benefit of predicting that some *after* constructions – namely those that receive  $\succ$  final interpretations and are thereby derived using MAX – do license NPIs (exceptionally). This prediction, coupled with the assumptions that the truth conditions of *after* modifiers require a  $\succ$  final interpretation and that NPIs are licensed only in unambiguously (Strawson-)downward-entailing contexts, accurately predicts that the only *after* constructions that license NPIs are those that are overtly modified. However, this account cannot predict the observation in (3) (from Krifka, 2010b) that only ‘long’ modifiers license NPIs under *after*; it incorrectly predicts that NPIs are just as acceptable under *right after* as they are under *long after*.

In his response to Condoravdi (2010), Krifka (2010b) proposes a different approach with what he touts as several advantages: it avoids likening multiple interpretations of *before* and *after* sentences to aspectual coercion, instead characterizing them as instances of semantic under-specification; it avoids the use of possible worlds in his semantics; and he eschews Strawson-entailment as a licensing condition for NPIs in favor of the one advocated for in Krifka 1995, in which NPIs come about via the use of a semantically weak proposition compared to its alternatives.

The latter represents a clear improvement over the B&C approach for its ability to predict

the (gradient) differences in acceptability of NPIs under *after* modifiers of various lengths, as illustrated in (3). But the elimination of possible worlds in the theory is of minimal advantages to those of us who endorse their use in general, and leaves us in a relatively unsatisfying position in terms of correctly characterizing the truth conditions of these sentences.

For Krifka, *before* and *after* are lexical converses or antonyms, associated with conversely but non-strictly ordered scales, as in (13).

- (13) a.  $\llbracket \text{before } B \rrbracket = \lambda t \neg \exists t' [t \leq t' \wedge \llbracket B \rrbracket (t')]$   
 b.  $\llbracket \text{after } B \rrbracket = \lambda t \exists t' [t \leq t' \wedge \llbracket B \rrbracket (t')]$

He highlights as a happy outcome the parallels between the meaning in (13a) and the degree-semantic meaning of Italian *prima* ('earlier') proposed in del Prete (2008). But that proposal came from idiosyncratic properties of *prima* in relation to English *before* and also in relation to Italian *dopo* ('after'). Krifka's analysis of both *before* and *after* as akin to *prima* incorrectly ignores these differences, and parallel differences in English between *before* and *after*, on the one hand, and *earlier* and *later*, on the other.

The relatively semantically underspecified entries in (13) also require a reliance on what Krifka refers to as 'pragmatics': the derivation of various implicatures to produce either the multiple meanings associated with some *before/after* constructions, or the relatively narrow meaning associated with them, depending on the context of utterance. For instance, the veridicality in (1a) (*Mozart died after he finished the Requiem*) comes from a conversational implicature that the EE (the *B* argument) is "reasonably probable" (p. 920), the result of what would otherwise be an "unmotivated restriction" (ibid.) that *B* is not true at any time before *A*. The non-strict ordering lexically encoded in both entries in (13) turns into a strict-ordering as the result of an implicature associated with the competition of both relations with *when*, which just requires overlap. And, finally, *before* and *after* compete with each other (p. 922), which generates, for each, an implicature that the other is not appropriate, which can optionally constrain the truth conditions as well.

However, there is little precedence for many of these implicatures when we turn to related phenomena. In degree semantics, it is largely assumed that the non-strict equative (*as Adj as*) competes with the strict comparative (*Adj-er than*), resulting in the former carrying an implicature that the strict comparison is not true (Rett, 2014). Analogously, this would predict that *when* – the semantically weaker of the relations – carry a Quantity implicature, not *before* or *after*. Similarly, if *before* and *after* are in fact antonyms – and further if *before* is the marked, negated version of *after*, as Krifka argues – the clear prediction is that only the marked antonym carries the Manner implicature, not that they are both symmetrically associated with an implicature (Rett, 2015b). The weakness of these pragmatic explanations, coupled with the relatively weak truth conditions imposed by (13), arguably over-generate the truth conditions in Table 1, predicting, for instance, that *before* sentences with stative or process EEs (e.g. *John met Mary before she was president*, (4a)) is ambiguous, like its culmination EE counterpart (5a).

### 3. A more general proposal

In what follows, I will characterize *before* and *after* in a way that closely parallels the semantic treatment of relations in other ordered domains: they will be context-sensitive, but more impor-

tantly they will be polar-sensitive, sensitive to the temporal ordering of their arguments. The result is an analysis that correctly predicts the truth conditions in Table 1; has the same capacity as the account in Condoravdi (2010) to address the veridicality and NPI asymmetries (for the same reasons); and involves only independently motivated components: 1) a treatment of *before* and *after* as antonyms, associated with scales with reverse orderings (and a scale-sensitive characterization of maximality); 2) two aspectual coercion mechanisms.

### 3.1. Order-sensitivity in the degree domain

Just as *before* and *after* relate two temporal intervals, antonymic comparatives like *taller* and *shorter* relate two degree intervals. The relations therefore have a lot in common; I will focus on those commonalities here and draw conclusions about the semantic treatment of temporal relations from the canonical treatment of comparatives.

But I will note that there are important morphosemantic differences between temporal relations like *before* and *after* and adjective-based comparatives like *earlier* and *later*. The morphological differences are relatively transparent; the semantic differences, highlighted in a very useful study in del Prete (2008), range from subtle to obvious (e.g. *earlier* and *later*, qua comparatives, both license NPIs in their embedded arguments, and so do not have an NPI asymmetry). So the goal of this section is to explain what sorts of lessons we can draw from comparatives – qua relations between strictly ordered arguments – without reducing the analysis of *before* and *after* to a degree-comparative-based analysis of *earlier* and *later* (cf. Krifka (2010b)).

#### 3.1.1. Cross-polar anomaly

Kennedy (1997) observes that, while comparatives can be formed with two distinct adjectives or parameters (14), these adjectives can differ with respect to dimension of measurement, but they cannot differ only with respect to polarity (15).

- (14) a. A is taller than B is wide.  
 b. A is more tall than B is wide.
- (15) a. \*A is taller than B is short.  
 b. \*The hose is shorter than the ladder is long.

The generalization he comes to (although see Büring, 2007) is that comparatives can only relate two arguments with the same ordering. This means that, even in a typical (one-adjective) clausal or even phrasal comparative, we can assume both arguments are ordered along the same scale. This is arguably intuitive, in an apples-to-oranges sense, and is what I will carry over to the treatment of *before* and *after*.

There is widespread agreement that positive antonyms (like *tall*) are associated with a scale consisting of points (modeled by  $(0, \infty]$ , the positive numbers) on a greater-than ( $>$ ) ordering. There are consequently two different ways of characterizing negative antonyms. One is to maintain that negative antonyms, too, range over positive numbers, but encode the opposite scale, less-than ( $<$ ; Option 1 in Table 3).

A second option is to characterize negative antonyms, too, as encoding the greater-than ordering, but to characterize them as instead ranging over points corresponding to the negative numbers (Option 2 in Table 3).

	<b>Option 1</b>	<b>Option 2</b>
<i>points</i>	positive $(0, \infty]$	negative $[-\infty, 0)$
<i>scale</i>	less than $<$	greater than $>$

Table 3: possible characterizations of negative antonyms

The latter is familiar from work on comparatives in Rullmann (1995) and Heim (2000), but translates less intuitively to the domain of times, in which the concept of negative times seems less intuitive. I will therefore characterize negative antonyms (like *short*, but also *before*) along the lines of Option 1: as associated with converse scales, but the same domain of time points. This, coupled with the restrictions illustrated by cross-polar anomaly, will require that we relativize the notion of maximality to the scale associated with the plural it ranges over.

### 3.1.2. Scale-sensitive maximality

As discussed in detail in Rullmann (1995), there seem to be a wide variety of natural-language phenomena whose semantic interpretations require a notion of maximality. These include comparatives, like *John swam faster than Bill could run*, and degree questions, like *How many children does John have?* It is in principle possible to model the semantics of these constructions without a maximality operator – by i.e. characterizing the comparative as a strict subset relation, or using the Gricean notion of Quantity to capture strong exhaustivity, even in embedded questions – but Rullmann argues that the use of a maximality operator can account for semantic subtleties (e.g. negative island effects in comparatives) that the others can't. He proposes something like (16), for a plurality  $X$  in any domain.

$$(16) \quad \text{MAX}(X) = \iota x[x \in X \wedge \forall x' \in X[x' \leq x]]$$

Of course, a maximality operator is just as necessary in comparatives with negative antonyms (like *John swam slower than Bill could run*) as it is in comparatives with positive antonyms. And, as Beck and Rullmann (1999) point out, it is needed for degree questions formed with upward-scalar predicates, like *How much money can a graduate student live on?* This notion of maximality takes for granted that both positive and negative antonyms order along the 'greater than' scale (with negative numbers), as in Option 2 in Table 3 above. To adapt it for Option 1, we will need to define maximality in a scale-sensitive way. This is done in (17), where  $R$  ranges over the relations  $>, <$ .

$$(17) \quad \text{MAX}(X_R) = \iota x[x \in X \wedge \forall x' \in X[(x' \neq x) \rightarrow x R x']]$$

This amounts to a local (i.e. non-propositional) implementation of the widely recognized need for a notion of 'maximal informativity' as opposed to maximality writ large (Dayal, 1996; Beck and Rullmann, 1999; von Stechow et al., 2014); it outputs  $n + 1$  over  $n - 1$  for a 'greater than' scale and  $n - 1$  over  $n + 1$  for a 'less than' scale.

### 3.2. Aspectual coercion

A final piece of the puzzle here is the importing of two independently motivated aspectual coercion mechanisms (de Swart, 1998): one for non-culminating processes (like states and activities), and another for culminated processes (like achievements and accomplishments).

The first is inchoative coercion, an interpretive option for states and activities. It captures the observation that processes can denote an interval corresponding to their entire runtime, as in (18a), or a single point corresponding to the onset or initial point of that interval, as in (18b) (interpreted as meaning that the start of Amy’s surprise happened when Betty paid the check).

- (18) a. Amy was surprised. *stative*  
 b. Amy was surprised when Betty paid the check. *inchoative*

Inchoative coercion is formalized in (19) (Dölling, 2014); several languages, like Russian, employ inchoative aspect markers (e.g. *-sja*) to derive the meaning in (18b) (Hamburger, 1984; Haspelmath, 1993).

- (19) **inchoative coercion:** If  $e$  is a process with duration  $T$ ,  $e$  can denote  $T$  or  $\text{GLB}(T)$ , where  $\text{GLB}(T) = \iota t[t \in T \wedge \forall t' \in T[t \leq t']]$

The second is completive coercion, a known interpretive option for accomplishments and achievements (i.e. telic eventualities or culminated processes; Dölling (2014)). In (20a), the accomplishment denotes an extended temporal interval; in (20b), it denotes a single temporal point associated with the telos, or culmination, of the accomplishment.

- (20) a. Jane climbed the mountain. *accomplishment*  
 b. Jane climbed the mountain at seven o’clock sharp. *semelfactive*

- (21) **completive coercion:** If  $e$  is a culmination with duration  $T$ ,  $e$  can denote  $T$  or  $\text{LUB}(T)$ , where  $\text{LUB}(T) = \iota t[t \in T \wedge \forall t' \in T[t \geq t']]$

I will assume that these coercion mechanisms, like any polysemy, can apply optionally to their arguments whenever context or aspectual marking allows. This means that, although the data in Table 1 are presented assuming that the main clause eventuality is punctual, the theory presented below predicts a broader range of truth conditions (process ME  $\times$  process EE; process ME  $\times$  culmination EE; culmination ME  $\times$  culmination EE; culmination ME  $\times$  process EE), including optional aspectual variation for each argument. I have reason to believe these broader predictions are correct, but do not have the space to present them in detail here.

### 3.3. The formal analysis

The formal analysis of temporal relations I propose combines these assumptions: the assumption (from cross-polar anomaly) that antonyms are associated with reverse orderings; the assumption that maximality operators are scale-sensitive; and the assumption that inchoative coercion and completive coercion can apply optionally to processes and culminations, respectively. Specifically, I propose that *before* and *after* encode reverse relations between some point in their matrix argument (here,  $A$ ) and the most informative point in their embedded argument (here,  $B$ ). In this sense, they are lexical converses, or antonyms, just as they are in Beaver and Condoravdi (2003), Condoravdi (2010), and Krifka (2010b).

- (22) a.  $\llbracket A \text{ before } B \rrbracket = \exists t \in A[t \prec \text{MAX}(B_{\prec})]$   
 b.  $\llbracket A \text{ after } B \rrbracket = \exists t \in A[t \succ \text{MAX}(B_{\succ})]$

In (22a), *before* relates some point in the matrix eventuality  $A$  to the maximal (on the ‘before’ ordering) time associated with the embedded eventuality  $B$ . And in (22b), *after* relates some point in the matrix eventuality  $A$  to the maximal (on the ‘after’ ordering) time associated

with the embedded eventuality  $B$ . A metalinguistic constraint – identical to the one needed to account for cross-polar anomaly – requires that the matrix and embedded temporal interval arguments both be ordered on the same scale as the one associated with the relation in order for the sentences to be defined.

### 3.3.1. The truth conditions

Direct application of the definitions in (22) account for four of the six possible interpretations of *before* and *after* sentences; incorporation of the two types of aspectual coercion account for the other two. I will begin by discussing the two ambiguous examples.

	process EE	example	derivation	culmination EE	example	derivation
$A$ before $B$	$\prec$ initial	(4a)	(22a)	$\prec$ initial $\prec$ final	(5a)	(22a) (20)+(22a)
$A$ after $B$	$\succ$ initial $\succ$ final	(5b)	(18)+(22b) (22b)	$\succ$ final	(4b)	(22b)

Table 4: predicting the readings of *before* and *after*

The sentence in (5a) is repeated in (23). Assume that John met Mary at 3pm, and that Mary climbed the mountain from 1pm to 4pm (i.e. that  $B = [1\text{pm}, 4\text{pm}]$ ). The  $\prec$  initial interpretation of this sentence comes from a direct application of (22a): the meeting time, 3pm, is related (on the precedence scale) to the earliest time in the EE interval (the time that imposes the strongest restriction on the temporal precedence relation). In this context, that reading is false (23a).

- (23) John met Mary before she climbed the mountain.
- a.  $= 3\text{pm} \prec \text{MAX}([1\text{pm}, 4\text{pm}]_{\prec})$   
 $= 3\text{pm} \prec 1\text{pm}$   $\prec$  initial: false
  - b.  $= 3\text{pm} \prec \text{MAX}(\text{completive}([1\text{pm}, 4\text{pm}]_{\prec}))$   
 $= 3\text{pm} \prec \text{MAX}([4\text{pm}]_{\prec})$   
 $= 3\text{pm} \prec 4\text{pm}$   $\prec$  final: true

The  $\prec$  final interpretation of this sentence comes from an application of completive coercion, appropriate because the embedded eventuality is a culmination. The temporal interval is first coerced into a singleton point representing its telos, and then MAX is trivially applied (in accordance with (22a)). The prediction is that the sentence is true iff 3pm indeed precedes 4pm, which it does.

The sentence in (5b) is repeated in (24). Assume that John met Mary in 2022, and that Mary was president from 2021-2028 (i.e. that  $B = [2021 - 2028]$ ). The  $\succ$  final interpretation of this sentence comes from the application of the maximality operator, which picks out the latest time in the EE interval, i.e. temporal endpoint. The result is a false reading of the sentence that claims that 2022 comes after 2028.

- (24) John met Mary after she was president.
- a.  $= 2022 \succ \text{MAX}([2021 - 2028]_{\succ})$   
 $= 2022 \succ 2028$   $\succ$  final: false
  - b.  $= 2022 \succ \text{MAX}(\text{inchoative}([2021-2028]_{\succ}))$   
 $= 2022 \succ \text{MAX}([2021]_{\succ})$

= 2022  $\succ$  2021

$\succ$  initial: true

The  $\succ$  initial interpretation of this sentence comes from an application of inchoative coercion, appropriate because the embedded eventuality is a process. The temporal interval is first coerced into a singleton point representing its temporal onset, and then MAX is trivially applied (in accordance with (22b)). The prediction is that the sentence is true iff 2022 comes after 2021, which it does.

Crucially, this account correctly predicts that two of the logically possible eight readings are unavailable. *Before* constructions with embedded processes cannot receive a  $\prec$  final interpretation, because that interpretation would require an aspectual coercion mechanism from a state to its endpoint, which is unattested. This asymmetry results from the facts that *before* privileges the greatest lower bound of the embedded interval and that processes cannot override this reading by coercing into a least upper bound point. And *after* constructions with embedded culminations cannot receive a  $\succ$  initial interpretation, because that would require an aspectual coercion mechanism from a telic eventuality to its beginning point, which is unattested. This asymmetry results from the facts that *after* privileges the least upper bound of the embedded interval and that culminations cannot override this reading by coercing into a greatest upper bound point.

And finally, the fact that this account relies on aspectual coercion mechanisms for two of the six available readings correctly predicts that languages like Russian, Serbo-Croatian, and Tagalog can disambiguate between the two readings for (5a) and (5b) using aspectual markers (like inchoative or completive markers, or perfective, etc.). This capability seems encouraging for an explanation of the fact that some languages differ in whether their *before* and *after* are in fact ambiguous in the ways English ones are, but I cannot account for such cross-linguistic variation here.

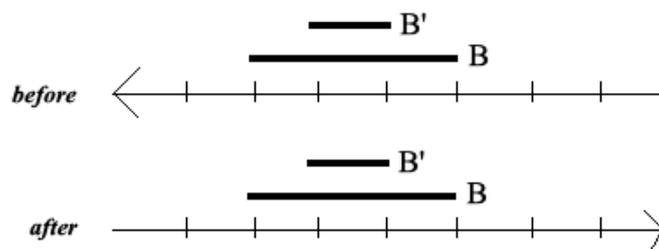
### 3.3.2. The universal asymmetries

In general, the present proposal accounts for the veridicality and NPI asymmetries just as Condoravdi (2010) does. Specifically, I will assume a branching theory of time in which past times can be evaluated in the actual world, but future times are necessarily modal. This, assuming intensionalized versions of *before* and *after* in (22), accounts for the observed veridicality asymmetry. It also, incidentally, accounts for the universality of the veridicality asymmetry: because the indeterminateness of the future is the same for every language, *before* is necessarily intensional, and its embedded eventuality is therefore not necessarily actualized.<sup>5</sup>

The current proposal also predicts that NPIs are licensed (given a Strawson-entailment characterization of NPI-licensing, von Stechow 1999) in *before* and *after* constructions but only under certain conditions.

As Figure 1 illustrates, aside from instances of completion coercion, whenever  $B'$  is a subset of  $B$ ,  $A \prec \text{MAX}(B_{\prec})$  entails  $A \prec \text{MAX}(B'_{\prec})$  (i.e. if some point in  $A$  precedes the initial point of  $B$ , it necessarily precedes the initial point of  $B'$ ). But this entailment relation does not hold if the point in  $A$  is related to the final points of  $B$  and  $B'$ . So the account predicts that *before*

<sup>5</sup>While Heinämäki (1974) argues that veridicality is conditioned by syntactic phenomena like NPIs and ellipsis, this doesn't ring true to me, nor is this claim endorsed by further analyses. Specifically, I can get a veridical reading for *I left the country before anything happened*, and a non-veridical reading for *John left before Bill*.

Figure 1: the monotonicity of *before* and *after*

constructions with process EEs, like (25a), license NPIs, and that *before* constructions with culmination EEs, like (25b), license NPIs, but only when they necessarily receive a  $\prec$  initial interpretation.<sup>6</sup>

- (25) a. John met Mary before she ever was president.  $\prec$  initial  
 b. John met Mary before she ever climbed the mountain.  $\prec$  initial or  $\prec$  final

And this indeed seems to be the case; (25b) cannot be used to accurately describe a situation in which John met Mary halfway up the mountain, i.e. after she summited the mountain. It remains to be seen whether this new observation is cross-linguistically robust (in other words, whether the universal NPI asymmetry reported above includes this nuance involving  $\prec$  final interpretations of *before* constructions).

I'll just highlight that this prediction is novel to the present analysis precisely because it is the first to explicitly account for initial/final ambiguities in *before* constructions. As indicated by the parallel scenario in Figure 1 – and knowing that *after* constructions are also initial/final ambiguous – my analysis predicts that *after* can license NPIs, but only in sentences that unambiguously receive  $\succ$  final interpretations. Figure 1 illustrates this downward-entailing property: whenever  $B'$  is a subset of  $B$ ,  $A \succ \text{MAX}_{\succ}(B)$  entails  $A \succ \text{MAX}_{\succ}(B')$  (i.e. if some point in  $A$  succeeds the final point of  $B$ , it necessarily succeeds the final point of  $B'$ ). But this entailment relation does not hold if the point in  $A$  is related to the initial points of  $B$  and  $B'$ , and therefore we can conclude that *after* constructions that are ambiguous between  $\succ$  initial and  $\succ$  final interpretations cannot license NPIs.

This is precisely the prediction made by the account in Condoravdi (2010).<sup>7</sup> However, in contrast to the analogous prediction about (25), this seems to be incorrect. Both constructions in (26) are characterized as ungrammatical, regardless of their interpretation.

- (26) a. \*John met Mary after she ever climbed the mountain.  $\succ$  final  
 b. \*John met Mary after she ever was president.  $\succ$  initial or  $\succ$  final

This is an unsatisfactory state of affairs, but it is in keeping with those made in Condoravdi (2010). And while Krifka (2010b) proposes an alternative account that makes better predictions regarding NPIs (and especially exceptional NPI-licensing), his account is lacking in a number

<sup>6</sup>It is typical to test NPI *any* in these embedded clauses, but I have chosen to use *ever* because *any* necessitates a verbal object, which has hard-to-control-for implications for the aspectual class of the embedded clause.

<sup>7</sup>I also inherit from her, in the absence of any other theoretical modification, the prediction that all modifiers of *after* equally license NPIs, which does not seem to be the case, as demonstrated in (3) (Krifka, 2010b).

of other areas. So I will have to leave this issue unsettled for now.

#### 4. Conclusion and discussion

Previous accounts of the semantics of *before* and *after* constructions have focused on the veridicality and NPI asymmetries between the two relations (Beaver and Condoravdi, 2003; Condoravdi, 2010; Krifka, 2010b). The assumptions in these papers regarding the truth conditions of the constructions have been unclear and incomplete. I have provided a little more scrutiny of the possible interpretations of *before* and *after* constructions, in English and other languages, and come to several novel empirical conclusions: the NPI and veridicality asymmetries are universal, and so should not be attributable to language-specific idiosyncrasies; there is no reason to think that these constructions bias the initial point of the embedded clause in any way (cf. Table 1); and *after* constructions are not unique in their ambiguity. I've also suggested that languages differ in the ambiguity of their *before* and *after* constructions, in a way that arguably tracks independent differences in aspectual marking and the interpretation of grammatical aspect. So at least some component of an analysis of the truth conditions of *before* and *after* should be appropriately aspectually sensitive.

The account presented here draws from the degree-semantics literature; in particular, from what we know about antonymic relations between ordered intervals in the degree domain. I adopt from that phenomenon (and from Kennedy, 1997) the assumption that both interval arguments of a relation must have the same ordering as the relation itself, and that any notion of maximality used to interpret these relations must be sensitive to the scale ordering of that interval.

The result is an account in which *before* privileges the initial point of its embedded argument and in which *after* privileges the final point of its embedded argument. When these definitions are combined with two known types of aspectual coercion – inchoative coercion for processes (states or activities), and completive coercion for culminations (achievements or accomplishments) – the account predicts all and only the six available interpretations of *before* and *after* constructions illustrated in Table 1. It also predicts that languages which employ aspect to overtly mark these aspectual coercion processes, like Serbo-Croatian (11), have correspondingly unambiguous *before* and *after* constructions.

Incidentally, as argued in Rett (2015a), this combination of scale-sensitive maximality and pragmatic processes that can make certain endpoints contextually salient predicts that relations between intervals are similarly ambiguous across strictly ordered domains. Specifically, it predicts the fact that positive comparatives like *faster* are generally unambiguously  $> \max$ , while negative comparatives like *slower* are generally ambiguous between a  $< \min$  and a  $< \max$  interpretation, as illustrated in the famous Lucinda sentences (Rullmann, 1995). The context-sensitivity of the coercion mechanisms – and their cross-domain counterparts – can also account for the observation that these ambiguities vary contextually (Beck, 2012).

- (27) a. Lucinda is driving faster than is allowed on this highway.  $> \max$   
 b. Lucinda is driving slower than is allowed on this highway.  $< \min$  or  $< \max$

As I argue in Rett (2015a), the parallel also carries over to locative prepositions, which can be thought of as relations between strictly-ordered spatial intervals, or vectors. The positive relation *above* in (28a) can only be interpreted as placing the bird above the top, or canopy, of

the tree; but the negative relation *under* in (28b) is ambiguous between a reading in which the bird is flying somewhere under the canopy (the < top interpretation) or, more darkly, buried in the ground beneath the roots of the tree (the < bottom interpretation).

- (28) a. The bird is above the tree. > top  
 b. The bird is under the tree. < top or < bottom

While the parallels are not absolute (recall that there is no NPI asymmetry in comparatives), I believe the truth-conditional similarities are similar enough to constitute evidence for a characterization of all antonyms as encoding scale reversals; all maximality operators as being scale-sensitive; and for there to be some degree of optionality, in strictly ordered domains, about whether the interval as a whole is available for a relation, or just a particular salient endpoint.

There are several adjacent phenomena I haven't had the opportunity to discuss. Condoravdi (2010) and Krifka (2010b) spend an admirable amount of time extending their treatments of *before* and *after* to the temporal relations *since* and *until*, which seem similarly antonymic. (*Since* sentences, like *after* sentences, situate the runtime their matrix eventuality after the runtime of their embedded eventuality, while *until* does the opposite.) I am optimistic that the general points of the present proposal extend naturally to *since* and *until* constructions, as do Condoravdi's and Krifka's accounts, but have not examined them in detail. Finally, I have left unaddressed at least one other intriguing cross-linguistic difference between *before* and *after*: *before* constructions, but not *after* constructions, can be formed with spurious negation, or expletive negation, in languages like German (Krifka, 2010a) and Mandarin (Lin, 2016).

## References

- Anscombe, E. (1964). Before and after. *The Philosophical Review* 73, 3–24.
- Beaver, D. and C. Condoravdi (2003). A uniform analysis of *before* and *after*. In *Proceedings of SALT XIII*, pp. 37–57. CLC Publications.
- Beck, S. (2012). Lucinda driving too fast again: the scalar properties of ambiguous *than*-clauses. *Journal of Semantics* 30, 1–63.
- Beck, S. and H. Rullmann (1999). A flexible approach to exhaustivity in questions. *Natural Language Semantics* 7, 249–298.
- Büring, D. (2007). Cross-polar nomalies. In *Proceedings of SALT 17*, pp. 37–52.
- Condoravdi, C. (2010). NPI licensing in temporal clauses. *Natural Language and Linguistic Theory* 28, 877–910.
- Dayal, V. (1996). *Locality in WH-quantification*. Kluwer.
- de Swart, H. (1998). Aspect shift and coercion. *Natural Language and Linguistic Theory* 16, 347–385.
- del Prete, F. (2008). A non-uniform semantic analysis of the Italian temporal connectives *prima* and *dopo*. *Natural Language Semantics* 16, 157–203.
- Dell, F. (1983). An aspectual distinction in Tagalog. *Oceanic Linguistics* 22/23, 175–206.
- Dölling, J. (2014). Aspectual coercion and eventuality structure. In *Aspects, phases, and arguments: topics in the semantics of verbs*, pp. 189–226. John Benjamins.
- von Stechow, K. (1999). NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16, 97–148.
- von Stechow, K., D. Fox, and S. Iatridou (2014). Definiteness as maximal informativeness. In L. Crnič and U. Sauerland (Eds.), *The art and craft of semantics: A festschrift for Irene*

- Heim*, Volume 1, pp. 165–174. MITWPL.
- Hamburger, H. (1984). Aktionsart and aspect in Russian. *Russian Linguistics* 8, 129–146.
- Haspelmath, M. (1993). More on the typology of inchoative/causative verb alternations. In B. Comrie and M. Polinsky (Eds.), *Causatives and transitivity*. John Benjamins.
- Heim, I. (2000). Degree operators and scope. In B. Jackson and T. Matthews (Eds.), *Proceedings of SALT X*, pp. 40–64. CLC Publications.
- Heinämäki, O. (1974). *Semantics of English temporal connectives*. PhD Thesis, UT Austin.
- Kennedy, C. (1997). *Projecting the adjective*. PhD Thesis, UCSC.
- Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic Analysis* 25, 209–257.
- Krifka, M. (2010a). How to interpret ‘expletive’ negation under *bevor* in German. In T. Haneforth and G. Fanselow (Eds.), *Language and Logos*, pp. 214–236. Akademie Verlag.
- Krifka, M. (2010b). *Before* and *after* without coercion: comment on the paper by Cleo Condoravdi. *Natural Language and Linguistic Theory* 28, 911–929.
- Lin, J.-w. (2016). Negation under *yiqian* ‘before’ in Mandarin Chinese and cross-linguistic variation in expletive negation. *Language and Linguistics* 7, 1–26.
- Moens, M. and M. Steedman (1988). Temporal ontology in natural language. *Computational Linguistics* 14, 15–28.
- Ogihara, T. (1995). Non-factual before and adverbs of quantification. In *Proceedings of Semantics and Linguistic Theory (SALT) V*. CLC Publications.
- Rett, J. (2014). Modified numerals and measure phrase equatives. *Journal of Semantics* 32, 425–475.
- Rett, J. (2015a). Antonymy in space and other strictly ordered domains. *Baltic International Yearbook of Cognition, Logic, and Communication* 10, 1–33.
- Rett, J. (2015b). *The semantics of evaluativity*. Oxford University Press.
- Rullmann, H. (1995). *Maximality in the semantics of wh-constructions*. Ph. D. thesis, University of Massachusetts, Amherst.
- Sharvit, Y. (2014). On the universal principles of tense embedding: the lesson from *before*. *Journal of Semantics* 31, 263–313.