eliminating EARLIEST:
a general semantics for before and after*

Jessica Rett
Sinn und Bedeutung 24
4 September, 2019

0 overview

• semantic differences between before and after have caused some to posit a null EARLIEST operator in the temporal domain (Beaver and Condoravdi, 2003; Condoravdi, 2010)
  – the operator is arguably theoretically unappealing (cf. LATEST; Krifka, 2010)
  – I present evidence from a variety of languages that EARLIEST is also empirically inadequate
• I propose a reformulation of the semantics of before and after in terms of a more general notion of maximal informativity (von Fintel et al., 2014)
• I draw parallels to similar arguments for eliminating MAX (Beck and Rullmann, 1999; Rett, 2015)

1 asymmetries between before and after

• we’ve known for a while that before and after are not inverses (Anscombe, 1964)
  1. before licenses NPIs, but after does not (Ogihara, 1995; Beaver and Condoravdi, 2003)
     (1) Amy arrived {before/*after} anyone left the party.
  2. after is veridical (entails the truth of its internal argument), but before is not (Heinämäki, 1974)
     (2) a. Mozart died before he finished the Requium. → Mozart finished the Requium.
        b. Mozart died after he finished the Requium. → Mozart finished the Requium.
• Anscombe (1964) attributes her observed asymmetries to a difference in quantificational force:

(3) a. \[ [A \text{ before } B] = 1 \text{ iff } \exists t \in A, \forall t' \in B[t < t'] \]
b. \[ [A \text{ after } B] = 1 \text{ iff } \exists t \in A, \exists t' \in B[t > t'] \]
  – this accounts for the veridicality and NPI asymmetries
  – but incorrectly predicts the truth conditions for before sentences (Beaver and Condoravdi, 2003)
    * e.g. that Squares had four sides long before David made a clean sweep of all of the gold medals in the Sydney Olympics is true in worlds in which David has never won a gold medal\(^1\)
    * and, impossibly, that Cleo left exactly 5 seconds before David sang is true for each time in the temporal extension of David’s singing

---

*I have benefited from discussion with Daniel Altshuler, Cleo Condoravdi, Sam Cumming, Gabe Greenberg, Jeff King, Nathan Klinedinst, Maziar Toorsarvardani, Yael Sharvit, and 2017 audiences at the UConn Logic Colloquium and MIT LingLunch. I am also indebted to Nicoletta Loccionì 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 Kudliy; Isabel Lin; Nicoletta Loccionì; 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 Committee on Research.

\(^1\)I don’t love this argument; we’ve already observed that before is non-veridical. But B&C cite the claim that we can’t analyze the non-veridicality of before sentences like the above as an underspecification (Landman, 1991; Valencia et al., 1992; Ogihara, 1995). They instead analyze it as a three-way ambiguity between veridical before, non-veridical before, and “the non-committal type of the non-veridical reading” (p47–8). I think this is an incorrect conclusion, especially in light of the cross-linguistic data.
• in response, Beaver and Condoravdi (2003) and Condoravdi (2010) (henceforth ‘B&C’) analyze before and after as follows (for times t, temporal intervals T, and \( \subseteq \) the temporal part relation):

\[
\text{EARLIEST}(T) = ut[t \in T \land \forall t'[t' \in T \rightarrow t < t']]
\]

- the NPI asymmetry is predicted because (4) characterizes a (Strawson-)downward-entailing environment in the B argument of before but not after
- the veridicality asymmetry is predicted with the adoption of a branching model of time, combined with the intensionalizing of EARLIEST
- (4) also predicts that, when the embedded B interval is temporally extended, before and after relate some time t in the A interval to the initial point of B.

2 effects of aspectual class on before and after

• this prediction – that before and after relate some time t in the A interval to the initial point of B – is empirically incorrect (Heinämäki, 1974)
• while B&C were aware of this, and addressed it, the empirical complications are broader than they had anticipated, and drastically undermotivate EARLIEST

2.1 known effects of aspectual class

• when after embeds a state, A is related to either the initial or final point of B, depending on context

\[
\text{[}_A \text{ John met Mary] before [}_B \text{ she was president].} \quad < \text{ initial}
\]

\[
\text{[}_A \text{ John met Mary] after [}_B \text{ she was president].}
\]

a. ...which is to say, after her inauguration, once she had settled into the role. \quad > \text{ initial}

b. ...which is to say, after she had retired from office. \quad > \text{ final}

– note: while > final logically entails > initial, speakers report that this is a true ambiguity, i.e. that the two readings are model-theoretically distinct
– B&C relate this ambiguity to inchoative coercion, and derive the non-inchoative > final reading by making EARLIEST optional

* open question 1: why isn’t EARLIEST optional for before (cf. the NPI prediction)?
* open question 2: what, then, ensures that these > final interpretations are still veridical?

• when before and after embed an accomplishment, they generally relate some point in the A interval to the final point in the B interval; it’s been claimed there is no ambiguity here\(^4\)

\[
\text{[}_A \text{ John met Mary] before [}_B \text{ she climbed the mountain].} \quad < \text{ final}
\]

\[
\text{[}_A \text{ John met Mary] after [}_B \text{ she climbed the mountain].} \quad > \text{ final}
\]

\(^2\)Where \( ut[t \in T \land \exists t'[t' \in T \land t' < t] \land \forall t''[t'' \in T \land \exists t'''[t''' \in T \land t''' < t''] \rightarrow t \not\subseteq t''] \)

\(^3\)Specifically, from Condoravdi 2010:892:

(i) \[ \text{[before X]}_A \text{ w} = \lambda t < \text{EARLIEST}_{alt_c(w, t, X)} X \] if defined

\[ \text{alt}_c(w, t, X) = \begin{cases} \{ w \} \text{ if } \exists t' \in T \text{ at}(w, t', X) \\ \{ w' \in \text{rep}_c(w, t) | \exists t' \in T \text{ at}(w', t', X) \} \end{cases} \] otherwise

"When X is not instantiated in the world of evaluation – the second case of [(i)] – the alternative worlds are selected from among the historical alternatives of the world of evaluation relative to a given time."

\(^4\)I use ‘accomplishment’ (Dowty, 1979) to refer to ‘culminated process’ (Moens and Steedman, 1988).
Heinämaki accounts for (8)–(9) by defining *before* and *after* with respect to the reference time of the embedded interval, stipulating that reference times are final points for accomplishments and initial points for other classes.

Condoravdi 2010:887 accounts for (8)–(9) by redefining *earliest* to operate on the output of *itop*, a function from an interval to its telos or right bound *rb* (if it has one).

\[
\text{itop}(T) = \begin{cases} 
\text{undefined if } T = \emptyset; \\
\{t \mid \exists t' \in T : rb(t') = t\} \text{ otherwise}
\end{cases}
\]

She therefore predicts that *before* and *after* both relate some time in *A* to the earliest or initial point of *B*, unless the embedded eventuality is non-homogenous (i.e. associated with a telos).

### 2.2 novel data & cross-linguistic variation

- but (8) is ambiguous; it also has a reading in which John met Mary before she started climbing

\[(11) \quad [\text{A John met Mary} \text{ before [B she climbed (to the top of) the mountain]}].\]

a. ...which is to say, they met on the bus ride to the mountain path. < initial

b. ...which is to say, they met halfway up the mountain. < final

- this means that Heinämaki’s reference time and Condoravdi’s *itop* accounts are too restrictive

- furthermore, the difference between the interpretations in (11) are conditioned by aspectual marking in some other languages

- in e.g. Russian and Serbo-Croatian, this difference is conditioned by the imperfective (< initial) and the perfective (< final)

\[(12) \quad \begin{array}{ll}
a. & \text{Vanja povstreˇ cal Maˇ su do togo, kak krasil dom.} \\
& \text{V met M before that how paintedIMPF house} \\
& \text{‘Vanja met Masha before he was painting the house.’} \text{ < initial} \\
b. & \text{Vanja povstreˇ cal Maˇ su do togo, kak porkasil dom.} \\
& \text{V met M before that how paintedPFV house} \\
& \text{‘Vanja met Masha before he painted the house.’} \text{ < final} \\
\end{array}\]

\[(13) \quad \begin{array}{ll}
a. & \text{Mary je srela Johna pre nego ?to se} \\
& \text{Mary-NOM is-PRES-3FS met-PP-3FS John-ACC before than PTCL REFL} \\
& \text{peo na vrh planine.} \\
& \text{climb-IMP-3MS on top-ACC mountain-GEN.} \\
& \text{‘Mary met John before he climbed to the top of the mountain.’} \text{ < initial} \\
b. & \text{Mary je srela Johna pre nego ?to se} \\
& \text{Mary-NOM is-PRES-3FS met-PP-3FS John-ACC before than PTCL REFL} \\
& \text{popeo na vrh planine.} \\
& \text{climb-PP-3MS on top-ACC mountain-GEN.} \\
& \text{‘Mary met John before he climbed to the top of the mountain.’} \text{ > final} \\
\end{array}\]

- in Tagalog, this difference is conditioned by the neutral, non-culminating perfective (< initial) and the culminating perfective (ability-and-involuntary-action, AIA, < final; Dell 1983)

\[(14) \quad \begin{array}{ll}
a. & \text{Um-alis siya bago niya w<in>alis-an ang-sahig.} \\
& \text{AV.PFV.NEUT-leave SUBJ.3sg before NON.SUBJ.3sg PFV.NEUT-sweep-LV SUBJ-floor} \\
& \text{‘She left before he swept the floor.’} \text{ < initial} \\
b. & \text{Um-alis siya bago niya na-walis-an ang-sahig.} \\
& \text{AV.PFV.NEUT-leave SUBJ.3sg before NON.SUBJ.3sg PFV.AIA-sweep-LV SUBJ-floor} \\
& \text{‘She left before he swept the floor.’} \text{ < final} \\
\end{array}\]
• a quick note on the broader typological study: I surveyed 17 languages (including English) from seven different language families

  – each language displayed the veridicality asymmetry between before and after originally reported for English
  – each language either didn’t license NPIs under temporal prepositions (Greek & Spanish) or only licensed NPIs under before (not after)
  – the after+state ambiguity exemplified in (7) is also universal (from Estonian)

(15) Sue tegutses peale Peteri abielus olemist.
Sue took-action after Peter married being
‘Sue took action after Peter was married.’ > initial or final

• to summarize:

<table>
<thead>
<tr>
<th></th>
<th>B is a state</th>
<th>example(s)</th>
<th>B is an accomplishment</th>
<th>example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A before B</td>
<td>&lt; initial</td>
<td>(6)</td>
<td>&lt; initial</td>
<td>(11)</td>
</tr>
<tr>
<td></td>
<td>&lt; final</td>
<td>(12)–(14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A after  B</td>
<td>&gt; initial</td>
<td>(7)</td>
<td>&gt; final</td>
<td>(9)</td>
</tr>
<tr>
<td></td>
<td>&gt; final</td>
<td>(15)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: truth conditions of before and after

• these data suggest that teloi vary in their salience based on considerations of context (as in (7) and (11)) and grammar (as in (12)–(14) and (15)).

• thus the meanings of before/after sentences justify neither an earliest operator (because initial points are not privileged) nor strict aspectual-class stipulations (because classes do not behave uniformly)

3 a more general proposal

• some desiderata:

  – an account of the NPI asymmetry, e.g. (1)
  – an account of the veridicality asymmetry, e.g. (2)
  – a potential ambiguity (contextual or grammatical) for before+accomplishments or after+states
  – no potential ambiguity for before+states or after+accomplishments

3.1 the formal analysis

• the analysis in a nutshell: before and after are better characterized as order- and context-sensitive, relating A to the most informative salient bound of B

  1. before and after reference the most informative salient bound of the embedded interval;
  2. the most informative salient bound is defined in an order-sensitive way;
  3. two known types of aspectual coercion apply to states and accomplishments

---

5 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

6 Although see del Prete 2008 for an interesting discussion of some complications in Italian.

7 I tested activities as well; they behaved like states for all intents and purposes.
1. *before* and *after* both encode a relation between some point in their *A* argument and the most informative point in their *B* argument

(16) a. \[A \text{ before } B] = \exists t \in A [t < !<(B)] \]
    b. \[A \text{ after } B] = \exists t \in A [t > !<(B)] \]

   • this approach is inspired by treatment of the degree relations *more* and *less* (Kennedy 2001; §3.3)
   • specifically, it assumes that the interval arguments to temporal relations are ordered in the same direction, specified by the relation itself (< for *before*, > for *after*), cf. *A is taller than B is short*
   • two comparative examples, for *A* (who is 5ft tall) and *B* (who is 6ft tall)

(17) a. ‘*A is more tall than B*’: relates (0,5ft] to (0,6ft] using the greater than relation
    b. ‘*A is less tall than B*’: relates [5ft,0) to [6ft,0) using the less than relation

   • two examples, for activities *A* from 2pm-5pm and *B* from 6pm-10pm

(18) a. ‘*A before B*’: relates [2pm, 5pm] to [6pm,10pm] using the precedence relation
    b. ‘*A after B*’: relates [5pm, 2pm] to [10pm,6pm] using the succession relation

2. the order-sensitive relation !\(o\) is adapted from an order-sensitive definition of (degree) maximality (Dayal, 1996; Beck and Rullmann, 1999; Heim, 2000)

(19) !\(o\)(T) = \text{def}
     For all points \(t, t'\) (type \(i\)); intervals \(T\) (type \(i,e\)); and relations \(R\) (type \(i,e,i,e\)) with orderings \(o\): \(!o(T) = t \in T[\forall t' \in T[\{t^r : t^r \in R\} \subseteq \{t^s : t^s \in R\}]]\)

   • this relation returns the most informative bound given the ordering: an initial endpoint (a glb) when applied to an interval argument of *before*, and a final endpoint (a lub) when applied to an interval argument of *after*

   • there’s plenty of discussion in the literature about the naturalness and empirical adequacy of such an approach:
      – Fox and Hackl (2007) argue that the Maximal Informativity approach is independently motivated with the additional assumption that natural language treats all intervals as dense
      – von Fintel et al. (2014) generalize an order-sensitive \(\text{max}\) to propositions: a “uniquely maximal object is the one that creates the most informative true proposition” (p167)

3. aspectual coercion (Smith, 1997; de Swart, 1998): two coercion mechanisms independently associated with states and accomplishments, respectively

(20) **inchoative coercion**: If \(e\) is a state with duration \(T\), \(e\) can denote \(T\) or \(\text{glb}(T)\)

   • inchoative coercion “shifts a predicate of occurrences to a predicate of boundaries which are the beginning of such an occurrence” (Dölling, 2014, 214)

(21) a. Amy was dumbfounded.
    b. Amy was dumbfounded when Betty paid the check.

   • some languages, like Russian, have inchoative aspect markers (e.g. -sja, Hamburger 1984; Haspelmath 1993)

(22) **completive coercion**: If \(e\) is an accomplishment with duration \(T\), \(e\) can denote \(T\) or \(\text{lub}(T)\)

   • completive coercion shifts an extended, telic event to its telos (Dölling, 2014)

(23) a. Jane climbed the mountain.
    b. Jane summited the climbed at seven-o-clock sharp.
3.2 the results

3.2.1 the truth conditions

- in the absence of any aspectual coercion, the definitions in (16) and (19) account for several of the available readings;
- the two forms of aspectual coercion in (20) and (22) account for the other two readings

<table>
<thead>
<tr>
<th>A before B</th>
<th>B is a state</th>
<th>example(s)</th>
<th>derivation</th>
<th>A after B</th>
<th>B is an accomplishment</th>
<th>example(s)</th>
<th>derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; initial</td>
<td>(6)</td>
<td>(16-a)</td>
<td></td>
<td>&gt; initial</td>
<td>(11)</td>
<td>(16-a)</td>
<td></td>
</tr>
<tr>
<td>&gt; final</td>
<td>(15)</td>
<td>(16-b)</td>
<td></td>
<td>&gt; final</td>
<td>(9)</td>
<td>(16-b)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7)</td>
<td>(16-b)+(20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: the account the readings of before and after

- two representative examples, assuming:  
  - John met Mary at 3pm
    - Mary climbed the mountain from 12pm to 4pm

(24) [[John met Mary before she climbed the mountain]] < initial; (11-a)
    a. = 3pm < ![12pm, 4pm]
    b. = 3pm < 12pm ✓

(25) [[John met Mary before she climbed the mountain]] < final; (11-b)
    a. = 3pm < ![completive([12pm, 4pm])]
    b. = 3pm < ![4pm]
    c. = 3pm < 4pm ✓

- two others assuming:
  - John met Mary in 2022
  - Mary was president from 2021-2028

(26) [[John met Mary after she was president]] > final; (7-b)
    a. = 2022 > ![2028, 2021]]
    b. = 2022 > 2028 ✓

(27) [[John met Mary after she was president]] > initial; (7-a)
    a. = 2022 > ![inchoative([2028, 2021])]
    b. = 2022 > ![2021]
    c. = 2022 > 2021 ✓

- crucially, two unavailable readings:
  - < final for before embedding states:
    would require an aspectual coercion from a state to its endpoint, which is unattested (cf. (21))
  - > initial for after embedding accomplishments:
    would require coercion from an accomplishment to its starting point, also unattested (cf. (23))

- in languages like Russian, Serbo-Croatian, and Tagalog, instead of aspectual coercion, the readings in Tables 1 & 2 are derived unambiguously with aspectual markers

---

8I employ standard interval notation, where an open interval (a, b) is {x: a < x < b}; a closed interval [a, b] is {a ≤ x ≤ b}; and partially closed intervals (a, b] and [a, b) are {x: a < x ≤ b} and {x: a ≤ x < b}, respectively.
3.2.2 the universal asymmetries

- the veridicality asymmetry:
  - the definitions of before and after in (16) can be intensionalized to account for the veridicality asymmetry, given a branching model of time
  - (this follows the B&C explanation of the veridicality asymmetry quite closely)

- the NPI asymmetry:
  - typical tests for downward-entailingness in temporal sentences incorrectly vary individuals in the temporal argument, as opposed to the temporal argument itself (from Linebarger, 1987):  
    (28) John got sick after he ate a green vegetable. → John got sick after he ate kale.
  - the right sort of test instead varies the subsets of the actual argument of the element in question
    (29) context: Mary was president from 2021-2028; Mary solved climate change in 2022.
    Mary won the Nobel Peace Prize after she was president.
    → Mary won the Nobel Peace Prize after she solved climate change.
  - but note that the entailment in (29) only works for the > final interpretation of the first sentence
  - Linebarger (1987): long after is NPI-licensing, too, but only in > final cases (Condoravdi, 2010)
    (30) a. She persisted long after she had any hope at all of succeeding.
    b. He kept writing novels long after he had any reason to believe they would sell.
  - (16) dutifully characterizes both before and after as (Strawson) downward-entailing with respect to the interval arguments...
  - ...but only in cases without inchoative coercion (i.e. when the > initial reading is unavailable)
  - the NPI asymmetry is therefore predicted by the analysis – as is exceptional NPI licensing in after clauses – provided that we assume NPIs are only licensed in unambiguously DE environments

3.3 generalizing to other domains

- to account for e.g. *Jane drove faster than Mary didn’t, Rullmann (1995) proposed an order-insensitive maximality operator:
  (31) Rullmann’s MAX
    Let $D$ be a set of degrees ordered by the relation $\leq$, then
    \[ \max(D) = \{ d \in D \land \forall d' \in D\, [d' \leq d] \} \]

- a few concerns:
  - this absolute formulation doesn’t generalize well to negative antonyms (Heim, 2007)
  - nor upward-monotonic questions (Beck and Rullmann, 1999)

- an alternative proposal: the effects of MAX are instead the result of a more general principle requiring maximal informativity (Dayal, 1996; Beck and Rullmann, 1999; Heim, 2007; Fox and Hackl, 2007; von Fintel et al., 2014; Rett, 2015)

- Krifka’s (2010: 916) explanation, made explicitly with regard to before and after, is even more general: “We might also argue that EARLIEST is motivated by general cognitive principles, as “edges” are salient in perceptual experiences as in hearing and vision; it is well-known that there are neurons that are detectors of sudden changes.”

---

\(^9\)This is also a typical problem in DE tests of degree constructions (Rett, 2010).
• Rett (2015) argues that maximality is:
  (a) domain-general;
  (b) order-sensitive; and
  (c) only calculated on closed bounds

\[ \text{Economy of Bounds} \]
\[
\begin{align*}
[R](B_o)(A_o) & \iff \exists x \in A [x \geq_o \max(\text{bound}(x', B_o) \land x' \in B)], \text{ where} \\
\text{Bound}(x, X_o) & = \exists x \in X_o [\forall x'[X_o(x') \rightarrow x >_o x']] \\
\text{For some relation } R \text{ and intervals } A \text{ and } B, \text{ ‘} ARB \text{’ is true iff some member of } A \text{ exceeds (on the relevant scale) the greatest closed bound of } B. \]
\]

• this Economy of Bounds principle can be seen as a slightly more complex and domain-general version of (19) that can also account for the similar behavior of the spatial prepositions above and under and comparatives like faster and slower

\[ \text{a. Lucinda is driving faster than is allowed on this highway.} \quad \text{(Rullmann, 1995)} \]
\[ \text{b. Lucinda is driving slower than is allowed on this highway.} \]

\[ \text{German Context} \]
The highway’s speed laws impose a maximum (120km/h) and a minimum speed (70km/h).
\[ \text{a. the positive relation (33-a) is true if Lucinda is driving 130km/h (i.e., above the maximum)} \]
\[ \text{b. the negative relation (33-b) is true if she is driving 60km/h (i.e., below the minimum)} \]

\[ \text{California Context} \]
The maximum speed on the highway is 120km/h (there is no minimum).
\[ \text{a. the positive relation (33-a) is true if Lucinda is driving 130km/h (i.e., above the maximum)} \]
\[ \text{b. the negative relation (33-b) is true if Lucinda is driving 100km/h (i.e., below the maximum)} \]

4 summary & conclusions

• there is no sense in which before or after privilege the earliest bound of their argument(s)...

• ...which makes the null operator EARLIEST empirically unmotivated as well as theoretically unattractive

• the account proposed here:
  – borrows from the degree-semantic literature the idea that relations between ordered sets are calculated with an order-sensitive notion of ‘maximal informativity’ (!)
  – borrows from the aspectual literature the idea that some aktionsart classes are polysemous between an interval reading and a specific endpoint reading (the beginning point for states, the endpoint for accomplishments)
  – assumes that some languages use aspect marking to differentiate between these two readings

• this account predicts the observed veridicality and NPI asymmetries, and also correctly characterizes the truth conditions of before and after sentences...

• ...using (and only using) independently motivated pragmatic considerations of maximal informativeness and aspectual coercion

• the pragmatic generality of the notion of maximal informativity is borne out when we turn to other (ordered) domains, like degrees and vectors (Rett, 2015)
references