Was I speaking before I spoke?[*][†]

Stefan Kaufmann
University of Connecticut

Converging on Causal Ontology Analysis (COCOA)
October 12, 2022

Goals of this talk:

- Take a close look at the relationship between non-veridical (esp. "counterfactual") uses of *'before'*-sentences and counterfactual conditionals
- Characterize and compare the notions of "likelihood" involved in the interpretation of non-veridical *'before'*-sentences and counterfactuals
- Explore some cross-linguistic data to get an idea of the range of variability in the interpretation of expressions of temporal precedence

---

1 Some facts about clausal *'before'* [1]

1.1 Entailments of *'before'*- and *'after'*-sentences

> The clausal complement of *'before'* is not entailed by the whole sentence. In contrast, the complement of *'after'* is entailed.[2] The matrix clause is always entailed:
>
> (1)   a.   *'A before B'* $\Rightarrow B$ *'A before B'* $\Rightarrow A$
>   b.   *'B after A'* $\Rightarrow B$ *'B after A'* $\Rightarrow A$

'A before B' entails A but not B[3]

(2)   I left the party before I got sick.   (3)   Mozart died before he finished the Requiem.
       a. $\Rightarrow$ I left the party.   a. $\Rightarrow$ Mozart died.
       b. $\not\Rightarrow$ I got sick.   b. $\not\Rightarrow$ Mozart finished the Requiem.

'B after A' entails both A and B

(4)   I got sick after I left the party.   (5)   Mozart finished the Requiem after he died.
       a. $\Rightarrow$ I got sick.   a. $\Rightarrow$ Mozart finished the Requiem.
       b. $\Rightarrow$ I left the party.   b. $\Rightarrow$ Mozart died.

> ➼ The truth of B (jointly with A) is not necessary for the truth of *'A before B'*.
> **BUT** *something* about B is required when B is false:
>
> (6) $A, \neg B \not\Rightarrow$ *'A before B'*

The following are adapted from Beaver and Condoravdi (2003):

(7)   David never won a gold metal at anything, but he once at lots of ketchup.
       $\Rightarrow$ David ate lots of ketchup before he won all the gold medals in the Sydney Olympics.
       $\not\Rightarrow$ Squares had four sides long before David won all the gold medals in the Sydney Olympics.

> ➼ Under what circumstances is *'A before B'* true when B is false?

---

[1]In this talk I am not concerned with uses of *'before'* with DP complements.

[3]What is entailed is of course not merely the truth of the matrix clause, but its truth at a time whose location is restricted by the temporal clause. I leave that implicit.

1.2   (Non-)veridical uses of *'before'* and their modal implications

> *'A before B'* has three possible interpretations[4] depending on the context: *veridical*, *non-committal*, and *counterfactual*

Veridical: *B* is true. No relevant modal implications.

> (8)   I was lucky: My train left the airport before the bomb exploded.
> $\rightsquigarrow$ *There was an explosion.* $\not\rightsquigarrow$ *An explosion seemed likely.*

Non-committal: *B* may or may not be true. If it is not, it was "reasonably probable".[5]

> (9)   Having seen suspicious pieces of luggage, I left the airport before any of them exploded.
> $\rightsquigarrow$ *There may or may not have been an explosion.* $\rightsquigarrow$ *An explosion seemed likely.*

Counterfactual: *B* is false. But it was reasonably probable.

> (10)   Fortunately, the police defused the bomb before it exploded.
> $\rightsquigarrow$ *There was no explosion.* $\rightsquigarrow$ *An explosion seemed likely.*

1.3   Beaver and Condoravdi (2003)

> Roughly speaking: *'A before B'* is true iff the reference time of *A* precedes the time of the earliest *B*-eventuality. If *B* didn't happen, check when *B* (plausibly) *could have* happened.

Details (skippable)

Evaluation at world *w*:

- If there are *B*-eventualities at *w*, take the earliest one of those.
- If *B* is not instantiated at *w*, check alternative worlds:
    - Trace *w* back to the reference time t_A of *'A'* (when *'not A'* was still possible)
    - Look at those courses of events that were *reasonably probable* at w, t_A
    - Check whether *'B'* is consistent with those courses of events
    - If consistent, then *earliest* is defined at *w* (as the earliest occurrence of *'B'* among the courses of events that were reasonably probable at w, t_A)

(Non-)veridical implications as contextual entailments:

- The definedness of *earliest* is not a pragmatic presupposition.
  (I.e., need not be satisfied at all worlds inthe context set in order for the interpretation to succeed.)
- *Point-wise* evaluation at worlds in the context set: Those where *earliest* is undefined are discarded.
  (I.e., where *B* did not happen and was unlikely at the relevant time).
- ➡ *'before/after'*-clauses can be informative.
- Only those worlds survive the update at which *B* either happened or was *reasonably probable*.
  (Note that if *B* happened, it need not have been likely!)
- For consistent update, it is sufficient that there be *some* such worlds in the context set.

---

[4]See also Heinämäki (1972); Ogihara (1995); Beaver and Condoravdi (2003). Beaver and Condoravdi show that this context dependence can be accounted for with a uniform analysis of *'before'* that does not postulate an ambiguity, hence it is misleading to treat the three interpretations as different *readings* of *'before'*.

[5]The term "reasonably probable" was first used by Beaver and Condoravdi (2003).

➠ Veridical and non-veridical readings are *contextual entailments*.
  (I.e., they result from updaing particular types of contexts with a *'before/after'* sentence)

Most important for present purposes:

- "For the counterfactual reading, the input context has to entail that A's occurrence makes B's later occurrence impossible and that prior to A's occurrence there was a process that made B's occurrence at least reasonably probable." (Beaver & Condoravdi 2003:51)
- Facts at times later than the reference time of *'A'* are given up in the process of "rerunning history."
- This suggests a parallelism between "counterfactual" *'A before B'* and the corresponding counterfactual conditional *'If had not been A, would have been B'*.

  (11) a. The police defused the bomb before it exploded.
       b. If the police had not defused the bomb, it would have exploded.

Questions:

- What does "reasonably probable" mean?
- What is the relationship between (11a) and (11b)?
- Are posterior facts (later than the reference time of *A*) ignored in both?

---

2 Counterfactual conditionals

First hypothesis based on the preceding section:

- Counterfactual *'A before B'* and the counterfactual conditional *'if had been not-A, would have been B'* go hand in hand.
  I.e., whatever makes one true (false) also makes the other true (false).
- Likelihood at the relevant past time is crucially involved in both.
- Spoiler: This is wrong.

2.1 *'Before'*-sentences and counterfactuals

Scenario 1

You are on the bus from Kyoto to Tokyo. You get off at Nagoya; the bus travels on. The next day you hear that the bus had an accident and everyone on board was injured or killed. Consider now the following sentences:

(12) If I hadn't gotten off the bus, I would have been injured.   *[true]*

(13) I got off the bus before I got injured.   *[false]*

- (13) is false because the accident *was not foreseeable* at the time I got off – i.e., likelihood.
- But (12) is true even though the accident was unlikely.

➠ Main difference:

  – The truth of the counterfactual depends on "posterior" facts (after the reference time of *'A'*).
  – The *'before'*-sentence does not.

Posterior facts

Naïve (and wrong) strategy for counterfactuals *'If had been A, would have been B'*[6]:

- "Rerun history" from a time *t* at which *'A'* was still a live possibility
- Restrict the possible continuations to those in which *'A'* occurs and that are most likely from the perspective of *t*
- The counterfactual is true iff all of those courses of events lead to *'B'*.
- ➡ Facts at times later than the reference time of *'A'* are given up in the process of "rerunning history."

Better strategy for counterfactuals: Pay attention to *causal (in)dependence*. Hold on to facts that are causally independent of the antecedent.

- *Bus example:* Whether the accident occurs is causally independent of my pre/absence on the bus.
- ➡ The accident was unlikely, but it did occur. This fact is *not* given up.

Counterfactuals *vs. 'before'*:

- Posterior facts that are causally independent of *'A'* are held constant in the interpretation of the counterfactual *'if had been ¬A, would have been B'*.
- Such posterior facts are given up in the interpretation of *'A before B'*.

NB. "Counterfactual *'before'*" is a misnomer.

---

3 *'Before'* again

Second hypothesis in contrast to the findings on counterfactuals:

- What matters for "counterfactual" *'A before B'* is whether *'B'* was likely at the time.
  (This wasn't right for counterfactuals, but may be right for *'before'*.)
- Consider *all and only* the facts up to reference time of *'A'*.
- Spoiler: This is still wrong.

3.1 Likelihood and *'before'*

Problem for the second hypothesis:

- Counterfactual *'A before B'* does not always imply that *'B'* is (was) likely at the time in question.
- Case in point: *A* may be likely, and *A* interferes with *B*.

---

(14) a. *[As meteors usually do / As scientists expected,]*
        the meteor burned up before it hit the ground.
    b. If the meteor had not burned up, it would have hit the ground.

- The context implies that the meteor could not possibly hit the ground. Yet (14a) is true.
- Notice that (14b) is also true. This is for a somewhat different reason, though (see below).

---

[6]Variants of this idea have informed many approaches in philosophy and linguistics (Downing, 1959; Adams, 1975; Ellis, 1978; Thomason and Gupta, 1981; Tedeschi, 1981; Dudman, 1994; Strawson, 1986; Bennett, 1988; Mellor, 1993; Edgington, 1995; Dahl, 1997, among others).

Similarly, (15a) is felicitous even though Bill's books were never likely to become overdue.

(15) a. *[As he always does / as people around here always do,]*
       Bill returned his books to the library before they became overdue.
    b. If Bill hadn't returned his books to the library, they would have become overdue.

The problem:

- *A* should be disregarded, lest it interfere with the truth of *B*.

  – the burning of the meteor
  – the returning of the books

- But *A* is also likely: if we were to bar it from the modal base, it creeps right back through the restriction to "reasonably probable" worlds.

One idea: write the restriction to non-*A* worlds into the truth conditions of *'A before B'*.
(In effect, make it more similar to the counterfactual.)
But there are reasons to be wary:

- It is less compositional. We couldn't evaluate *'before B'* by itself, without reference to *A*.
- It only applies in the veridical case. Consider:

  (16) John struck the match before it lit.

The solution: *disregard A* without forcing not-*A*.

- Intuitively, focus on a part of the world: the process that is/was (at the relevant time) bound to bring about the truth of the *'before'*-clause.
- Don't even think about things external to that process.
NB. This calls to mind B&C's mention of a "process". But that was rather vague. Can we say more?

A side note on counterfactuals

- The counterfactuals in (14b) and (15b) are also true.
- Conditional antecedents, where present, "win" over unmentioned conflicting facts.[7]
- In *'before'*-sentences, we want to check whether a given time lies "before *B*," regardless of what is asserted about that time.
- ➡ Our interpretation of *'before B'* should not make reference to (the negation of) *A*.

---

[7] See Hiddleston (2005) for a theory which takes this into account.

4 The progressive

4.1 Parallels between the Progressive and *'before'*-clauses

- Non-veridicality (aka the "Imperfective paradox" – Dowty 1977):

  The entailment from *'X was V-ing'* to *'X V-ed'* . . .

  . . . does not hold for accomplishments: . . . (although it does hold for activies):

  (17) a. Mary was walking to the store.   (18) a. Mary was walking.
       b. $\Rightarrow\!\!\!\!/$ Mary walked to the store.   b. $\Rightarrow$ Mary walked.

- No modality on the veridical reading (aka "The problem of non-interruptions" – Landman 1992)

  "If an accomplishment manages to get completed, it is unproblematic to assume (in retrospect) that the progressive is true during the development stage."   (p. 14)

  (19) [What is Mary doing in this photograph?]
          Well, I would never have believed it at the time, but in fact . . .
       a. she was crossing the Atlantic Ocean.
       b. this was taken before she crossed the Atlantic Ocean.

- Irrelevance of prior facts (aka "The problem of interruptions" – Landman 1992)[8]

  (20) Mary was crossing the street when the truck hit her.

  – Dowty (1977): *'X was V-ing'* is true if X V-ed in all *inertia worlds*.
  – Problem: The truck is present in all inertia worlds. You can't just edit it out.
  – Solution: Focus on *part of the world*
      (in Landman's proposal, the *event* of Mary's crossing).

  (21) *'Mary is crossing the street'* is true in *w* at *t* iff some process of crossing by Mary, *e*, is going on in *w* at *i* and in every inertia world for *w* and *e* at *i*, i.e., in every world where *e* is allowed to follow its normal course, there is an interval surrounding *t* where *'Mary cross the street'* is true.

  (22) In determining what chance a stage [of an event] has of continuing, we abstract away from facts about the world that are external to that stage. Thus we are not interested in the absolute chance that the stage has of continnuing, given all the facts about the world, but in its chance of continuing *solely on the basis of what is internal*, inherent to that stage.
       Thus, in the case of [Mary crossing the street], the truck is irrelevant.
                                              (emphasis in the original)

---

[8]Landman (1992) credits Vlach (1981) for this kind of counterexample, though his proposed solution differs somewhat from Vlach's. I am not going to dicuss the details here.

4.2 Some differences

For all the parallels, *'before'* and the Progressive are not always interchangeable.

Pretty clearly interchangeable:

(23) a. Mozart died before he finished the Requiem.
 b. Mozart was finishing the Requiem when he died.

(24) a. Mary was hit by a truck before she crossed the street.
 b. Mary was crossing the street when she was hit by a truck.

Pretty clearly non-interchangeable due to Aspect:

These cases suggest that the Progressive interacts with the aspectual properties of the complement more tightly than *'before'*. This can lead to infelicity for various reasons.

- Achievements: no activity phase; not always a clear preparatory process.

 (25) a. The police defused the bomb before it exploded.
  b. ??The bomb was exploding when the police defused it.
  c. ??The police defused the bomb when it was exploding.[9]

 (26) a. The meteor burned up before it hit the ground.
  b. ??The meteor was hitting the ground when it burned up.
  c. ??The meteor burned up when it was hitting the ground.

- Activities: The Progressive is only true during the event. No preparatory phase.

 (27) a. John had a drink before he slept.
  b. ??John was sleeping when he had a drink.

 (28) a. Mary swam before she ran.
  b. ??Mary was running when she swam.

- Statives: The Progressive doesn't work that well with statives; but when it can be used, the temporal interpretation doesn't correspond to the *'before'*-sentence.

 (29) a. Sue was tall before she was smart.
  b. ??Sue was being smart when she was tall.

 (30) a. Jill was in the park before she was in the gym.
  b. ??Jill was being in the gym when she was in the park.

Upshot:

- The Progressive comes constraints (related to aspectual properties) that *'before'* seems largely insensitive to.
- The two coincide most clearly with accomplishments. There, we can observe their numerous commonalities.
- The commonalities involve properties that *'before'* has with all aspectual classes; whereas the Progressive is bad with other aspectual classes for independent reasons.
- ➡ The analysis should abstract away from aspectual properties.

---

[9]Generally, both orders of these *'when'*-sentences stand or fall together. I will only list one in the examples below.

More non-interchangeability, for mysterious reasons:

- Landman (1992) argues (based on data credited to Igal Kvart) that the completion of the accomplishment can be too far-fetched for the Progressive.

- (31) is false if all Mary does is strike down two or three soldiers, then is killed herself.

  (31)   Mary was wiping out the Roman army.

- But it seems that *'before'* is not so sensitive. These judgments are tentative and rather subtle.

  (32)   a.   Mary was struck down before she wiped out the Roman army.
         b.   ??Mary was wiping out the Roman army when she was struck down.

  (33)   a.   Mary drowned before she crossed the Atlantic Ocean.
         b.   ??Mary was crossing the Atlantic Ocean when she drowned.

- I do not at this time have an account of this.

5 How to model all this

5.1 Temporal reasoning

- A *Dynamic Bayesian Network* is, essentially, a Bayesian Network with a temporal dimension (Dean and Kanazawa, 1989; Koller and Friedman, 2009).
- Simple example: three variables A, B, C, observed at three times, t_1, t_2, t_3; The dotted lines are *persistence links*, modeling the dependence of the value of a variable at t_{n+1} on its value at t_n.
- My notation: 'X?' means X's value is not set/observed; 'X' means X is true; '$\overline{X}$' means X is false.
- Diagonal arrows indicate interactions (prevention, enabling). B_3 depends on both A_2 and B_2.
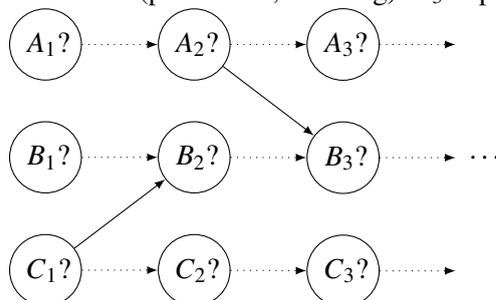
- Dependencies may be given by conditional probabilities (Kaufmann, 2005); ordering sources (Kaufmann, 2013); or structural equations (Schulz, 2011)
- A world is a complete setting of all variables to values.

5.2 The bus
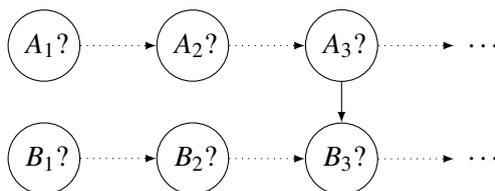
A? = Accident?; B? = Stay on bus?

- My bus ride is represented by a chain of positive *B*-nodes (different lengths at different worlds).
- Suppose at w I got off the bus at t_2; the accident happens at t_3. Thus:
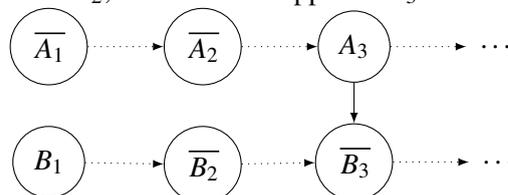
- For the counterfactual: Set B_2 to "true"; B_3 becomes true due to the persistence link. Leave A intact.
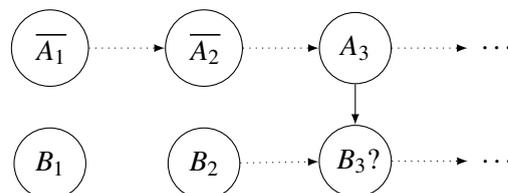
➥ If I hadn't gotten off the bus, I would have been injured. ✓
- For *'before'* (first stab): Undo all facts after the time in question.
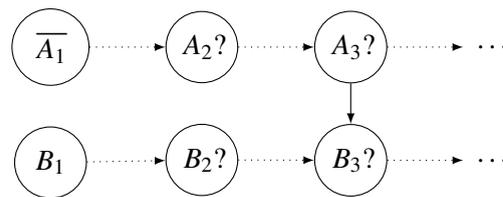
- Let the dependencies run their course. The accident is not likely to happen (A_2, A_3 false); the speaker is likely to stay on the bus (B_2, B_3 true).
 ➡ I got off the bus before I got injured. ✗

5.3 The truck

- Above the accident was unlikely to happen. But the truck was likely to hit Mary when she was crossing the street. More is needed to model that.
- Focus on "what is internal to the event" of Mary's crossing.
- I think of this as an extreme kind of network surgery: Cut all links into all slices of the event. But keep the persistence links (marginalize out the dependencies on variables external to it).
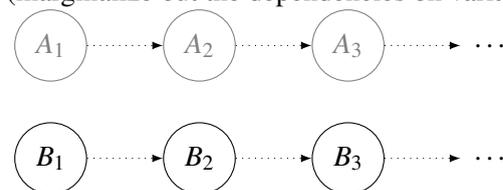
- ➡ Mary was hit by a truck before she reached the other side. ✓

5.4 Landman

- Compare Landman's (1992) notion of "Continuation branches" of events:
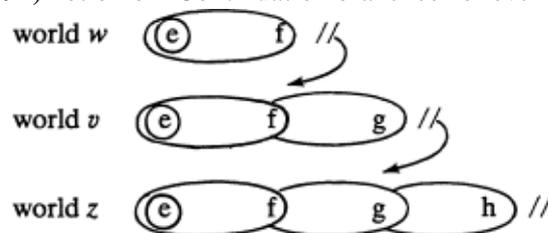
- Disregard worlds at which the event is interrupted; move to ones at which it continues, as long as that is "reasonable".
- In Landman's picture you go to more and more far-fetched worlds (until they become too far-fetched).
- In our picture, removing interruptions makes far-fetched worlds more llikely.

6 Conclusion

- Causal (in)dependence is at work in *'before'*, counterfactuals, and the Progressive.
- *'Before'* and counterfactuals are interpreted rather differently: Hold independent facts constant for counterfactuals, give them up for *'before'*. The Progressive is more closely tied to event structure than either of those.
- But the set of relevant variables is generally small (and likely subject to contextual and general cognitive factors).
- ➡ Causal structures are useful in the interpretation of all three (in different ways), and in abstracting away from event properties that matter only for the Progressive.

References

Adams, E. 1975. *The Logic of Conditionals*. Reidel.

Anscombe, E. 1964. Before and after. *The Philosophical Review*, 73.

Athanasiadou, A. and R. Dirven, editors. 1997. *On Conditionals Again*, volume 143 of *Amsterdam Studies in the Theory and History of Linguistic Science*. John Benjamins.

Beaver, D. and C. Condoravdi. 2003. A uniform analysis of *before* and *after*. In *Proceedings of SALT XIII*, pages 37–54.

Bennett, J. 1988. Farewell to the phlogiston theory of conditionals. *Mind*, 97(388):509–527.

Dahl, O. 1997. The relation between past time reference and counterfactuality: A new look. In Athanasiadou and Dirven (1997), pages 97–112.

Dean, T. and K. Kanazawa. 1989. A model for reasoning about persistence and causation. *Computational Intelligence*, 5(2):142–150.

Downing, P. B. 1959. Subjunctive conditionals, time order, and causation. *Proceedings of the Aristotelian Society*, 59:125–140.

Dowty, D. 1977. Towards a semantic analysis of verb aspect and the English 'imperfective' progressive. *Linguistics and Philosophy*, 1:45–77.

Dudman, V. H. 1994. On conditionals. *Journal of Philosophy*, 91:113–128.

Edgington, D. 1995. On conditionals. *Mind*, 104(414):235–329.

Ellis, B. 1978. A unified theory of conditionals. *Journal of Philosophical Logic*, 7:107–124.

Grandy, R. E. and R. Warner, editors. 1986. *Philosophical Grounds of Rationality*. Clarendon Press.

Harper, W. L., R. Stalnaker, and G. Pearce, editors. 1981. *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Reidel.

Heinämäki, O. 1972. Before. In *Proceedings of CLS 8*, pages 139–151.

Heinämäki, O. 1974. *Semantics of English Temporal Connectives*. PhD thesis, UT Austin.

Hiddleston, E. 2005. A causal theory of counterfactuals. *Noûs*, 39(4):632–657.

Kaufmann, S. 2005. Conditional predictions: A probabilistic account. *Linguistics and Philosophy*, 28(2):181–231.

Kaufmann, S. 2013. Causal premise semantics. *Cognitive Science*, 37:1136–1170.

Koller, D. and N. Friedman. 2009. *Probabilistic Graphical Models*. The MIT Press.

Landman, F. 1992. The progressive. *Natural Language Semantics*, 1:1–32.

Mellor, D. H. 1993. How to believe a conditional. *The Journal of Philosophy*, 90:233–248.

Ogihara, T. 1995. Non-factual *before* and adverbs of quantification. In *Proceedings of SALT V*, pages 273–291. CLC Publications, Cornell University.

Schulz, K. 2011. If you'd wiggled A, then B would've changed. *Synthese*, 179:239–251.

Strawson, P. F. 1986. 'If' and '⊃'. In Grandy and Warner (1986), pages 229–242.

Tedeschi, P. 1981. Some evidence for a branching-futures semantic model. In Tedeschi and Zaenen (1981), pages 239–270.

Tedeschi, P. and A. Zaenen, editors. 1981. *Tense and Aspect*, volume 14 of *Syntax and Semantics*. Academic Press.

Thomason, R. H. and A. Gupta. 1981. A theory of conditionals in the context of branching time. In Harper et al. (1981), pages 299–322.

Valencia, V. S., T. von der Wouden, and F. Zwarts. 1992. Polarity and the flow of time. In de Boer, A., J. de Jong, and R. Landeweerd, editors, *Language and Cognition*, pages 209–218. Uitgeverij Passage, Groningen.

Vlach, F. 1981. The semantics of the Progressive. In Tedeschi and Zaenen (1981).