

Before and After without Coercion:

Comment on the paper by Cleo Condoravdi

Manfred Krifka

The following contribution¹ was inspired by Cleo Condoravdi's article on NPI licensing in temporal clauses (Condoravdi, this volume). Condoravdi gives a coherent and comprehensive account of *before* which crucially involves coercion of propositions to the earliest or maximal times at which the propositions are true, and a modal component for non-factual interpretations. I argue for a non-modal, non-coercive analysis of clauses like [A *before* B] as 'A is the case when B has not been the case', triggering a conversational implicature that B will be the case later. I will discuss temporal operators involving measure phrases, like *three hours before*, and I will show that so-called "expletive" negation in corresponding clauses in German is, in fact, interpreted as negation.

1. *Before-* and *After-*: Condoravdi's analysis

The main purpose of Condoravdi's paper is to explain the occurrence of negative polarity items (NPIs) in the context of certain temporal prepositions, most prominently in the scope of *before*².

- (1) *The burglar left before anyone discovered the theft.*
- (2) *The burglar left before anyone noticed him.*
- (3) *The burglar left before anyone called the police.*

NPI licensing can be observed for the factual or veridical use of *before*-clauses as in (1), where we can assume that someone discovered the theft, and for their counterfactual use as in (2), where there is an entailment that nobody noticed the burglar. It also applies to the non-committal use as in (3), where it is not clear whether anyone called the police (cf. Heinämäki 1972, Ogihara 1995 for these distinctions). In contrast, *after*-clauses only have a factual use, and NPIs are not licensed in general, cf. (4), but only in particular cases as when modified by expressions denoting a long time, cf. in (5).

- (4) **The burglar left after anyone noticed him.*
- (5) *The burglar was scared long after there were any traces left.*

¹ Thanks to Arnim von Stechow, Daniel Büring and Sophie Repp for valuable criticisms and suggestions, and especially to Doris Penka for her comments in the preparation of this paper.

² Condoravdi only considers *before* subcategorizing clauses, but we find the same behavior with *before* heading DPs, as in *before any museum visit*, or heading gerunds, as in *before visiting any museum*. The temporal operators are considered prepositions and not complementizers because they allow for measure phrases as in *three hours before* in all their uses. Only the temporal uses of *before* and *after* will be considered here.

Condoravdi explains the occurrence of NPIs as a consequence of the interpretation of *before* proposed by Beaver & Condoravdi (2003). This interpretation creates a downward-entailing context, which in turn explains the occurrence of NPIs, cf. Ladusaw (1979).

The basic meaning of *before* is just the temporal order $<$ between time intervals; $t < t'$ holds iff the time t wholly precedes the time t' . Arguably, this holds for *before* as a preposition taking a nominal expression denoting a time, as in (6):

$$(6) \quad \llbracket \textit{before noon} \rrbracket = \lambda t[t < \textit{noon}]$$

If *before* heads a clause, which is assumed to refer to a set of times, the clause is coerced to a time. Condoravdi discusses two coercion operators, one reducing the set of times to the earliest time of the set, and one reducing it to the maximal time. Let us assume the following two definitions, where \sqsubseteq is the part relation for times.³

$$(7) \quad \text{EARLIEST}(T) = \iota t[t \in T \wedge \neg \exists t'[t' \in T \wedge t' < t] \wedge \forall t''[t'' \in T \wedge \neg \exists t'''[t''' \in T \wedge t''' < t''] \rightarrow t \sqsubseteq t'']]$$

$$(8) \quad \text{MAX}(T) = \iota t[t \in T \wedge \forall t'[t' \in T \rightarrow t' \sqsubseteq t]]$$

That is, $\text{EARLIEST}(T)$ is the smallest time in T not preceded by any other time in T , and $\text{MAX}(T)$ is the largest time in T provided that all other times in T are part of it. In both cases, this is only defined if there is such a time. In particular, $\text{MAX}(T)$ will be defined if T is cumulative, that is, whenever $t \in T$ and $t' \in T$, then the sum $t \sqcup t' \in T$.

In a slight variation from Condoravdi's representation that allows for a somewhat more perspicuous discussion, we take propositions to be functions from worlds into functions from times to truth values. We then have the basic meaning of *before* as in (9) and the interpretations as a clausal operator as in (10.a) after type-shift with EARLIEST or as in (b) after type-shift with MAX . The combination of a *before* clause with a main clause is interpreted conjunctively as in (11).

$$(9) \quad \llbracket \textit{before} \rrbracket = \lambda w \lambda t \lambda t' [t' < t] \quad \text{basic meaning}$$

$$(10) \quad \llbracket \textit{before B} \rrbracket \text{ a. } \lambda w [\lambda t \lambda t' [t' < t](\text{EARLIEST}(\llbracket B \rrbracket(w)))] \quad \text{type shift EARLIEST} \\ = \lambda w \lambda t [t < \text{EARLIEST}(\llbracket B \rrbracket(w))]$$

$$\text{ b. } \lambda w [\lambda t \lambda t' [t' < t](\text{MAX}(\llbracket B \rrbracket(w)))] \quad \text{type shift MAX} \\ = \lambda w \lambda t [t < \text{MAX}(\llbracket B \rrbracket(w))]$$

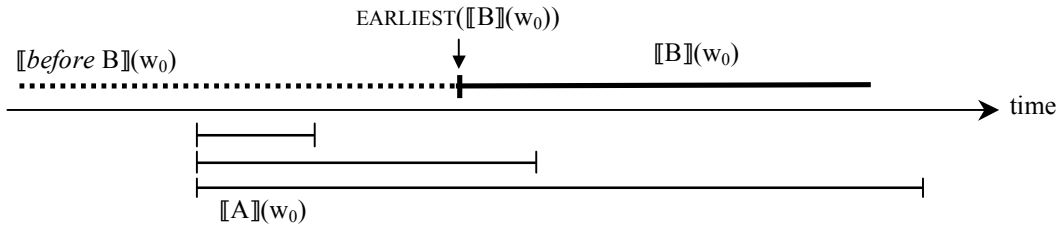
$$(11) \quad \llbracket A [\textit{before B}] \rrbracket = \llbracket [\textit{before B}] A \rrbracket = \lambda w \lambda t [\llbracket A \rrbracket(w)(t) \wedge \llbracket \textit{before B} \rrbracket(w)(t)]$$

It turns out that type shift with EARLIEST and MAX give us the same result provided that MAX is defined (which is the case if $\llbracket B \rrbracket$ is cumulative), and hence we just consider the (a) version here. If this sentence is true at world w_0 and time t_0 , then A must be true at w_0 and t_0 , and the first time at

³ Condoravdi gives no definition of the EARLIEST operator, and Beaver & Condoravdi (2003) define it as identifying for a set of times T the time t , $t \in T$, such that for all $t' \in T$ it holds that $t \leq t'$. It is not quite clear what $t \leq t'$ means for time intervals, and hence I suggest definition (7), partly following a suggestion by Doris Penka (pers. comm.). The operator MAX Condoravdi defines as mapping a set T that is cumulative to the sum of all T together with its greatest lower bound and smallest upper bound. The current definition is simpler and has similar effects. Furthermore, it corresponds to the definition of the definite article for domains with sum individuals, cf. Link (1983), which has been assumed independently as a coercion operator.

which B becomes true in w_0 is after t_0 . The following diagram illustrates this. We assume that B is a divisive predicate, that is, if $\llbracket B \rrbracket(w)(t)$ and $t' \sqsubseteq t$ then $\llbracket B \rrbracket(w)(t')$; hence EARLIEST singles out a minimal time at the beginning of $\llbracket B \rrbracket(w)$.

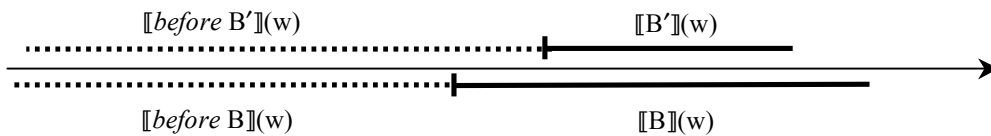
(12)



Notice that if A is atelic, then $\llbracket A \rrbracket(w_0)$ can also be true at larger times, even times that include the earliest point at which B is true, or that extend beyond the times at which B is true – as in *John lived in Paris before Mary did, and he continued to live there after she left*.

According to this theory, the *before*-clause creates a downward-entailing context, and hence should license NPIs. Assume that B' is a temporally more specific proposition, i.e. it holds that $\llbracket B' \rrbracket(w) \subseteq \llbracket B \rrbracket(w)$; it then follows that $\llbracket before B \rrbracket(w) \subseteq \llbracket before B' \rrbracket(w)$ and consequently it holds that if $\llbracket A \llbracket before B \rrbracket \rrbracket$ is true, then $\llbracket A \llbracket before B' \rrbracket \rrbracket$ is true as well:

(13)



The proviso here is that B' is true at some time, otherwise $\llbracket before B' \rrbracket$ is not defined, as the earliest time or maximal time of $\llbracket B' \rrbracket(w)$ is not defined if this is the empty set, as then the iota operator in the definitions of EARLIEST or MAX in (7) or (8) is not defined. Condoravdi argues that this is exactly what we need to explain a restriction of general downward-entailingness of *before* clauses, and hence a possible gap in the explanation why NPIs occur in such clauses. If downward entailment were general, the following inference would hold, but as a matter of fact it doesn't; it might be that someone other than the police discovered the theft.

(14) *The burglar left before someone discovered the theft.*
 \Rightarrow *The burglar left before the police discovered the theft.*

Condoravdi argues that the notion of downward-entailingness we need to explain NPI licensing is based on the notion of Strawson entailment (von Stechow 1999), where $\Phi \Rightarrow_{\text{Strawson}} \Psi$ iff for all worlds w such that the presuppositions of Φ and Ψ are satisfied, $\Phi(w) \Rightarrow \Psi(w)$. As the consequent clause in (14) presupposes that the police discovered the theft due to EARLIEST or MAX, we have to restrict our consideration to those worlds in which the police discovered the theft, and for these cases the entailment relation holds.

Condoravdi states that this restriction of downward-entailingness constitutes an essential advantage over previous accounts of *before*, in particular the quantificational proposal of Anscombe (1964), which can be implemented as in (15), as Anscombe's proposal does not create this presupposition:

(15) $\llbracket before B \rrbracket(w) = \lambda t \forall t' [\llbracket B \rrbracket(w)(t') \rightarrow t < t']$

It should be mentioned here that Anscombe could defend her account by requiring that in natural language, quantification is always over a non-empty domain, by which we would regain the presupposition. There is considerable evidence for that, including the way how universal quantification was understood in classical logic. If quantification is restricted to non-empty domains, then downward-entailingness would be restricted to those propositions B' for which $\llbracket B' \rrbracket(w)$ actually is defined.

The analysis outlined so far gives us the correct result for factual *before* clauses. For non-factual uses, $\text{EARLIEST}/\text{MAX}(\llbracket B \rrbracket(w))$ is undefined, as B fails to be true at any time in w , or is at least not guaranteed to be true at some time. Condoravdi argues, following Ogihara (1995), that in this case we switch to alternative courses of events that are considered “reasonably probable” in which B actually does become true. The sentence $[A \text{ [before } B]]$ then is true iff A holds for a time before the earliest time at which B becomes true in reasonable alternative worlds. This is illustrated in the following diagram, in which B does not become true in the real world but at reasonably probable alternatives. Under this analysis, non-factual *before*-clauses create downward-entailing contexts as well, and consequently they should endorse NPIs, which in fact they do.

(16)



As for *after*, Condoravdi (this volume), again based on Beaver & Condoravdi (2003), assumes the following derivations, which are exactly parallel to the derivations of *before* clauses:

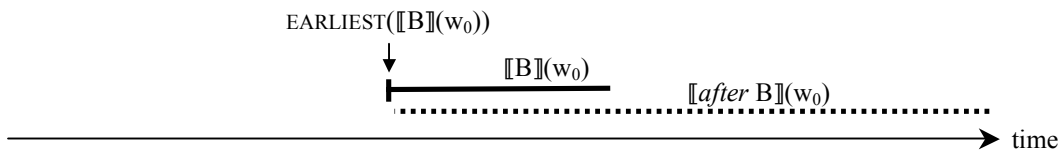
(17) $\llbracket \textit{after} \rrbracket = \lambda w \lambda t' \lambda t [t < t']$ basic meaning

(18) $\llbracket \textit{after } B \rrbracket$ a. $\lambda w [\lambda t' \lambda t [t < t'] (\text{EARLIEST}(\llbracket B \rrbracket(w)))]$ type shift EARLIEST
 $= \lambda w \lambda t [\text{EARLIEST}(\llbracket B \rrbracket(w)) < t]$
 b. $\lambda w [\lambda t' \lambda t [t < t'] (\text{MAX}(\llbracket B \rrbracket(w)))]$ type shift MAX
 $= \lambda w \lambda t [\text{MAX}(\llbracket B \rrbracket(w)) < t]$

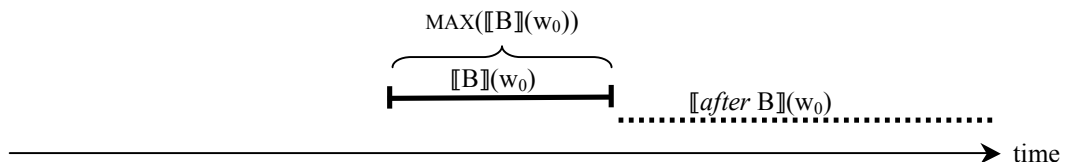
(19) $\llbracket [A \text{ [after } B]] \rrbracket = \llbracket [\textit{after } B] A \rrbracket = \lambda w \lambda t [\llbracket A \rrbracket(w)(t) \wedge \llbracket \textit{after } B \rrbracket(w)(t)]$

This gives us the two possible interpretations (20), (21) for $[\textit{after } B]$ with respect to a world w_0 :

(20)



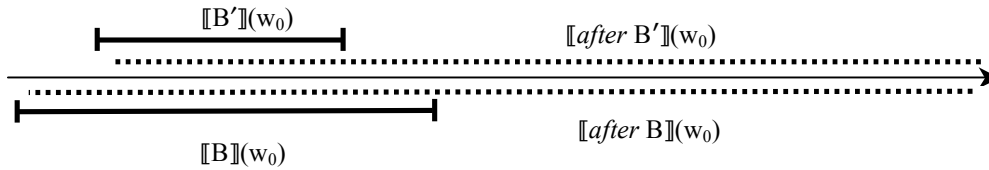
(21)



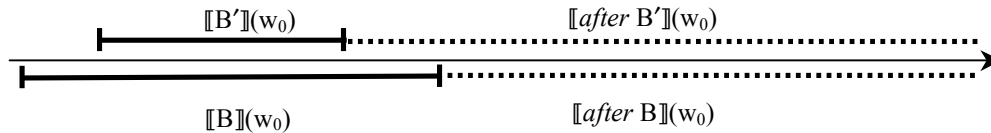
Coercing the B clause with EARLIEST does not create a downward-entailing context, cf. (22). That is, $\llbracket B' \rrbracket(w) \subseteq \llbracket B \rrbracket(w)$ does not guarantee that $\llbracket \textit{after } B \rrbracket(w) \subseteq \llbracket \textit{after } B' \rrbracket(w)$ – rather to the contrary. But

coercing with MAX, as in (23), yields just the mirror-image to (13), and hence provides a downward-entailing context.

(22)



(23)



This explains why in the case of cumulative B clauses, where MAX is defined, *after* clauses create a downward-entailing context, and hence allow for NPI items, as in (5) above.

(24) *The burglar was scared long after there were traces from his hands or boots left.*
 $\Rightarrow_{\text{Strawson}}$ *The burglar was scared long after there were traces from his boots left.*

Let me come to a preliminary assessment of Condoravdi’s account. It explains two aspects of the use of *before* and *after* elegantly: First, it captures the intuition that the two operators are essentially converses of each other, cf. their basic meanings (9) and (17). All apparent differences are due to coercion, in case the operators are applied to clauses. Second, the available coercion operators create the presupposition that there are times at which the temporal clauses are true – in the case of non-factual *before* this may be in other possible worlds. This explains the NPI-licensing behavior of *before* and *after*, as they can be shown to create downward-entailing contexts with respect to Strawson-entailment. Hence Condoravdi’s account constitutes clear progress over previous theories. However, there are a few points that I think are worth discussing.

Coercion is understood as an adaptation of a meaning following certain general rules to fit local requirements. Are MAX and EARLIEST such general rules? There is a good chance that they are. In particular, MAX – in the version proposed here – has been proposed for the definite article in languages that lack an article (cf. Partee 1987, Chierchia 1998). It also has been proposed for the aspectual domain to handle cases like *Mary jogged in thirty minutes*, where we understand *jogged* as referring to a maximal event of jogging (as in ‘did her jogging in thirty minutes’ – see Moens & Steedman (1988) and much subsequent literature). The other coercion operator, EARLIEST, has been invoked for inchoative coercion in cases like (25), cf. again Moens & Steedman (1988) and Fernald (1999).

(25) *Within an hour, Mary knew the answer.*

‘Within the hour following the reference time, Mary began to know the answer.’

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.

However, on closer look the case for EARLIEST is not particularly strong. (25) can also be interpreted as: Within the time interval of an hour following the reference time, there occurred a time t' at which Mary knew the answer. We can assume that under normal circumstances, if Mary knows the answer at time t' , then she also knows the answer at times that follow t' . For pragmatic reasons of

the same kind that have to be invoked for the inchoative analysis, the sentence will be adequate only if the first time t' at which Mary knew the answer occurs late in the interval of one hour, as otherwise alternative assertions giving shorter time spans would have been more informative. Hence we do not have to assume inchoative coercion in this case. And as for the argument concerning the perceptual saliency of edges, notice that this would also predict a coercion operator LATEST, which does not seem to exist; *before Mary lived in Paris* cannot be understood as *before Mary ceased to live in Paris*. It is unclear where this asymmetry should come from.

Another problem of EARLIEST as a coercion operator is that it is difficult to make it explicit, in contrast to MAX. While (26.a) sounds fine, (b) is at least a bit clumsy:

- (26) *Mary knew John before she knew Bill.*
 a. *Mary knew John before the time she knew Bill.*
 b. [?]*Mary knew John before the earliest / first time she knew Bill.*

Interestingly, neither paraphrase works well with the non-factual use of *before*. This comes as a surprise in Condoravdi's account, as EARLIEST and MAX apply in the derivation of this reading as well.

- (27) a. *Mozart died before he finished his requiem.*
 b. ^{???}*Mozart died before the (earliest) time he finished his requiem.*

As for NPI licensing, notice that in the non-factual case downward-entailingness is not restricted as in the factual case. For example, from (27.a) it follows that Mozart died before he finished his requiem for jazz combo, even though worlds in which Mozart actually finished a requiem for jazz combo are highly unlikely. This means that for the non-factual case we do not have to restrict the notion of downward-entailingness by using Strawson-entailment. In Condoravdi's analysis, we actually would restrict downward-entailingness in this way for the non-factual interpretation as well.

Condoravdi makes use of the motivation for NPIs as put forward in Krifka (1995). In that account, NPIs are expressions that come with alternatives, which result – in the case the NPI occurs in an assertion – in alternative assertions. For example, the assertion *Mary does not know any German person* evokes alternative assertions like *Mary does not know a Bavarian*, or *Mary does not know Harald Schmidt*. These alternative assertions are less restrictive than the assertion based on the NPI, with the understanding that the speaker did not base the assertion on other propositions because these propositions would be less informative. Now, there might be other reasons why a speaker did not make certain alternative assertions: While some of them would have been less informative, thus violating the maxime of quantity, others would have been lacking evidence, or be simply false, thus violating the maxime of quality. Thus, to say that *Mary went to Germany before she knew any German person* invokes a set of alternative propositions: First, propositions of the form 'Mary went to Germany before she knew x', where x ranges over the Germans that Mary actually did meet; these propositions are weaker than the one actually asserted, and they are not asserted, satisfying Quantity. Secondly, propositions of the type 'Mary went to Germany before she knew x', where x ranges over the Germans that Mary did not meet; these propositions are not asserted because they would violate Quality. Consequently, the NPI theory of Krifka (1995) does not require full downward-entailingness, and so we would not have to restrict downward-entailingness by invoking Strawson entailment.

Finally, I would like to draw attention to the fact that in Condoravdi's account, the A and B clauses are treated as lacking tense. There is only one tense operator over both clauses. This is because tense is assumed to bind the time argument of propositions, and we need that time argument to

combine the A and B clauses with *after* and *before*. There exist tenseless varieties, as in *before going to school*, but the examples discussed are finite and do have distinct tense operators.

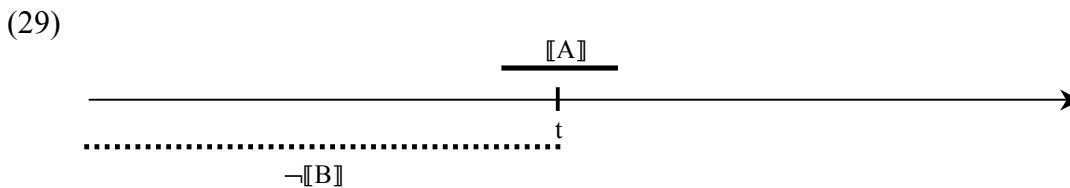
2. *Before* as temporal negation

Even though Condoravdi’s account is promising, and the points raised against it are certainly not conclusive, it is worthwhile to consider an alternative. It has the advantage that it deals with the non-factual readings in a simpler way, as it is non-intensional, in the sense that it only refers to times, and not to possible worlds. It also does not rely on type-shifting; in particular, it does not require the EARLIEST operator. As for MAX, I would like to keep it open whether this otherwise well-established operator is invoked. In any case, we do not have to assume it for the phenomena discussed here. The theory proposed here also makes substantial use of pragmatic implicatures generated by the meanings of *before* and *after*.

For *before* as a clausal preposition I would like to propose the lexical meaning (28). It negates that the *before* clause is true at or has been true before the time of interpretation of the A clause.⁴ Here, $t' \leq t$ expresses that the interval t' precedes or overlaps the interval t .

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

It will turn out that we do not have to recur to alternative possible worlds, hence propositions are simply represented as functions from times to truth values. The proposed interpretation can be illustrated as in (29). A is true at t , and can also be true before or after t . B cannot be true at the dotted line, but is allowed to be true after.



In addition, I will claim that the lexical meaning of *before* generates a conversational implicature that B will occur after t , to which I will turn shortly.

(28) is equivalent to Anscombe’s proposal, cf. (15), if the universal quantifier in that proposal is interpreted standardly, i.e. as not presupposing a non-empty domain. However, I prefer the form given here, as it makes it evident that B is not presupposed to be the case. It should be noted that, being a negation, *before* creates a downward-entailing context, and hence licenses NPIs, just as in Condoravdi’s proposal.

I should mention here that the current proposal has similarities to the treatment of Italian *prima* by del Prete (2008). In that paper, *prima* is analyzed as a comparative expressing a comparison of times, which are analyzed as degrees on a scale ordered by temporal precedence. Specifically, *prima* is analyzed as a combination of the meaning of *più* ‘more’ and *presto* ‘early’, hence expresses roughly the concept ‘earlier’. Comparatives in general are interpreted following Seuren

⁴ In Prior’s tense logic, $[A \textit{ before B}]$ is interpreted as $[A \wedge H\neg B]$, that is, A is the case, and it always has been the case that $\neg B$.

(1973), which in the current case results in the following meaning representation, where I use times instead of del Prete’s events, and disregard tense for simplicity.

- (30) *Gianni arrivò prima che arrivasse Lea.*
 Gianni arrive.PAST before than arrive.PAST.SUBJ Lea.
 ‘Gianni arrived before Lea arrived’

$$\begin{aligned} & \exists d[\neg[\text{Lea arrived d-early}] \wedge \text{Gianni arrived d-early}] \\ & \Leftrightarrow \exists d[\neg\exists t[\text{Lea arrived at } t \wedge t \text{ is d-early}] \wedge \exists t[\text{Gianni arrived at } t \wedge t \text{ is d-early}]] \\ & \Leftrightarrow \exists t[\neg\exists t'[\text{Lea arrives at } t' \wedge t' \leq t] \wedge \exists t'[\text{Gianni arrives at } t' \wedge t' \leq t]] \end{aligned}$$

The last inference follows from the monotonicity properties of degrees. While different in details, del Prete’s interpretation also contains a negation in the *prima* clause, which results in similar effects as will be discussed here. But there is little evidence that English *before* is a comparative, in contrast to Italian *prima*; for example, it does not occur with the comparative complementizer *than*, whereas *prima* licenses the comparative complementizer *che*. Also, del Prete points out that *dopo* ‘after’ is a different type of operator, whereas I will maintain that *before* and *after* are closely related – one is simply the negation of the other. But it is not the goal of this article to provide an in-depth comparison with the comparative approach to expressing the notion of ‘before’.

Back to our proposal. It gives a satisfying account for the non-factual reading, where B either does not become true at all, or it remains open whether B becomes true. We can illustrate this with the following example, where the tenses are fully interpreted; here *n* is the time of interpretation.

- (31) $\llbracket \text{Mozart finished his requiem} \rrbracket^n = \lambda t[t < n \wedge \text{Mozart finishes his requiem at } t]$
 $\llbracket \text{Mozart died} \rrbracket^n = \lambda t[t < n \wedge \text{Mozart dies at } t]$
 $\llbracket \text{Mozart died before he finished his requiem} \rrbracket^n$
 $= \lambda t[t < n \wedge \text{Mozart dies at } t \wedge \neg\exists t' [t' \leq t \wedge t' < n \wedge \text{Mozart finishes his requiem at } t']]$

If the resulting presupposition is asserted, the time variable *t* is existentially quantified. Notice that the dependent clause is not independently asserted, and the tense variable is bound by *before*.

In Condoravdi’s account of the non-factual interpretation of *before*, which is based on Ogihara (1995), we have to resort to other possible worlds that are deemed probable. In the case at hand, the account would entail that worlds in which Mozart finished his requiem were quite probable at the time before he died. This meaning component is absent in the current proposal. But we can derive it from general pragmatic rules, as a sentence [*A before B*], in its non-factual reading, is only informative in contexts where B is considered reasonably probable to be true at some time. If the common ground carries the information that B is highly unlikely at **any** time, then the statement that B is not true at **any time before a time at which A is true** is an unmotivated restriction. In more formal terms, if the a-priori likelihood of $\exists t[B(t)]$ is close to 0, then the a-priori likelihood of [*A before B*] is close to 1 (provided that A is true), and asserting such a proposition would violate the maxim of relevance, and hence the principle of cooperativity. Consequently, when uttering a sentence [*A before B*], the speaker creates the implicature that the a-priori probability that $\exists t[B(t)]$ is substantially greater than 0.

Notice that there is tendency for *before*-clauses that they are not just interpreted as likely, but rather as true. For example, on hearing out of the blue that Mr. Jones met Mrs. Smith before he met Mr. Smith, we readily infer that Mr. Jones actually did meet Mr. Smith. But this is a conversational implicature that can be cancelled:

(32) *Mr. Jones met Mrs. Smith before he met Mr. Smith – in fact, he never met him at all, because they divorced soon after.*

This tendency for a factive interpretation can be explained by the relation between the a-priori likelihoods of $\exists t[B(t)]$ and $[A \text{ before } B]$ that we have observed above. The informativity of $[A \text{ before } B]$ will be increased most if B in fact is true. Hence we can assume an implicature that $\exists t[B(t)]$ is not just likely, but in fact true.

This explains why sentences of the type $[A \text{ before } B]$ are often used in cases in which B is already established in the common ground. In Condoravdi's account, B is always presupposed by virtue of the MAX operator. In the present account, the B sentence is not literally presupposed (that is, *before*-clauses do not conventionally trigger a presupposition). But the peculiar pragmatics of $[A \text{ before } B]$ sentences is best satisfied if B actually is true.

If $[A \text{ before } B]$ states that B was not true before A became true, but implicates that B is highly likely or even true, then this is tantamount to saying that $[A \text{ before } B]$ implicates that B became true **after** A became true. It is useful for the upcoming discussion to make this implicature explicit, as in the following representation, where it is rendered in boldface.

(33) Assertion of $[A \text{ [before } B]] = \exists t[A(t) \wedge \neg \exists t'[t' \leq t \wedge B(t')]; \exists t'[\mathbf{B}(t')]$
 $\Leftrightarrow \exists t[A(t) \wedge \neg \exists t'[t' \leq t \wedge B(t')]; \exists t'[\neg t' \leq t \wedge \mathbf{B}(t')]$
 $\Leftrightarrow \exists t[A(t) \wedge \neg \exists t'[t' \leq t \wedge B(t')]; \exists t'[t < t' \wedge \mathbf{B}(t')]$

I have used the connector “;” which is intended to be understood like conjunction; I will discuss an important difference in section **Fehler! Verweisquelle konnte nicht gefunden werden.** It is important to keep in mind that the boldfaced part is not part of the lexical meaning of *before*, but an implicature that can be cancelled. Also, it should be mentioned that in cases in which the speaker knowledge may be incomplete, a weaker form of the implicature – that it is possible that B becomes true after t – may arise, which would cover the non-committal use of *before*.

In the following representation, the implicature is integrated into the meaning, as in (34).

(34) $\llbracket \text{before } B \rrbracket = \lambda t[\neg \exists t'[t' \leq t \wedge B(t')]; \exists t'[t < t' \wedge \mathbf{B}(t')]$

As the claim that B is true at some later time is just an implicature, and not an entailment, the current theory still predicts that *before* creates a downward-entailing context, if we take into account the lexical meaning only. If this appears unsatisfying, then recall that within the theory of NPIs in Krifka (1995) we do not require full downward-entailingness, as argued for above.

Turning now to *after*, we observe that this is just the negation of *before*, which satisfies the pretheoretical intuition that *before* and *after* are converses of each other.

(35) $\llbracket \text{after } B \rrbracket = \llbracket \text{not [before } B] \rrbracket = \lambda t \exists t'[t' \leq t \wedge \llbracket B \rrbracket(t')]$

As an example, consider the following derivation:

(36) $\llbracket \text{after [Mary knew John]} \rrbracket^n$
 $= \lambda t \exists t'[t' \leq t \wedge [t' < n \wedge \text{Mary knows John at } t']]$
 $\llbracket \text{Mary knew Bill [after [she knew John]]} \rrbracket^n$
 $= \lambda t [t \leq n \wedge \text{Mary knows Bill at } t \wedge$
 $\quad \exists t'[t' \leq t \wedge [t' < n \wedge \text{Mary knows John at } t']]]$

This is true at times t before the speech time n where Mary knows Bill at t , and that are either overlapped or preceded by a time t' that is before n such that Mary knows John at t' . This gives us nearly the same truth conditions as Condoravdi's analysis, where the B clause is coerced by EARLIEST. The only difference is that Condoravdi would have the condition $t' < t$, where we have $t' \leq t$. As a matter of fact, it seems that \leq is the correct choice, given examples like (37), where the two times appear to be contemporaneous.

(37) *The light went off after she tripped the switch.*

The reason why *after* is often understood as involving a temporal relation $<$, and not \leq , is presumably due to the fact that it is in pragmatic competition with *when*, which indicates a relation of temporal overlap. This means that *after* clauses are enriched by an implicature, rendered again in bold-face, where \circ is the temporal overlap relation:

(38) assertion of $[A \text{ [after B]}]$, with implicature due to competition with *when*:
 $\exists t[[A](t) \wedge \exists t'[t' \leq t \wedge [B](t')]]$; $\neg \exists t[[A](t) \wedge \exists t'[t' \circ t \wedge [B](t')]]$

For accomplishment or achievement clauses, as in *Mr. Jones met Mr. Smith after he met Mrs. Smith*, this enforces A to be true after B ceased to be true. For atelic clauses, especially stative clauses as in *Mr. Jones knew Mr. Smith after he knew Mrs. Smith*, the implicature does not arise because if such clauses are true at a time t , they continue to be true at later times, thus necessarily leading to temporal overlap. But there is still the competition with another temporal operator, *while*, which indicates temporal inclusion. This will ensure that in $[A \text{ after B}]$ the time at which B is true starts earlier than the time at which A is true.

(39) assertion of $[A \text{ [after B]}]$, with implicature due to competition with *while*:
 $\exists t[[A](t) \wedge \exists t'[t' \leq t \wedge [B](t')]]$; $\neg \forall t[[A](t) \rightarrow \exists t'[t \sqsubseteq t' \wedge [B](t')]]$

For completeness, it should be mentioned that $[A \text{ after B}]$ implicates that A is not true before the time the B clause is true. For example, *Mary was happy after John arrived* comes with the implicature that Mary was not happy before John arrived. Similarly, $[A \text{ before B}]$ implicates that A is not true after B. All these implicatures can be cancelled.

(40) Assertion of $[A \text{ [after B]}]$, with implicature due to competition with *before*:
 $[A \text{ [after B]}] = \exists t[[A](t) \wedge \exists t'[t' \leq t \wedge [B](t')]]$; $\neg \exists t[[A](t) \wedge \neg \exists t'[t' \leq t \wedge [B](t')]]$

(41) Assertion of $[A \text{ [before B]}]$, with implicature due to competition with *after*:
 $[A \text{ [before B]}] = \exists t[[A](t) \wedge \neg \exists t'[t' \leq t \wedge [B](t')]]$; $\neg \exists t[[A](t) \wedge \exists t'[t' \leq t \wedge [B](t')]]$

As with *before* clauses, the B clause often is presupposed. But presupposing B in *after*-clauses does not have the same dramatic effect as in *before*-clauses, where it switches from the non-factual reading to the factual reading. *After*-clauses are always factual. The reason is that B can be derived from $[A \text{ after B}]$ even if B is not presupposed.

It should also be obvious that *after*-clauses, in the current representation, do not create a downward-entailing context, and hence do not license NPIs. There is no negation or other entailment-reversing operator in the representation (35). This leaves us with the problem of how to explain those cases in which *after*-clauses in fact do license NPIs.

In this section, I have tried to show that the non-modal, non-coercive analysis of Anscombe (1964), when supplied with some pragmatic enrichments, actually looks quite good if compared with Condoravdi's approach. In the following sections, I will turn to a number of additional phenomena.

3. *Before* and *after* with temporal measure phrases

There is a use of *before* and *after* in which these temporal operators are combined with temporal measure phrases. Let us consider a case with *before*:

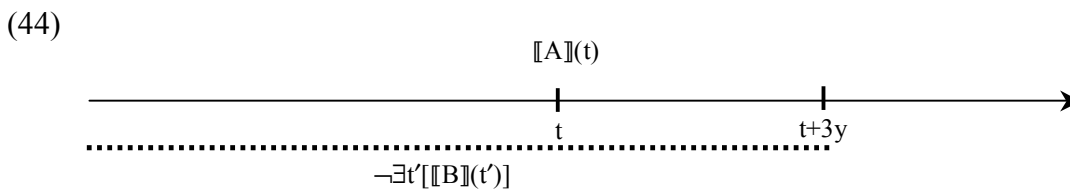
(42) *Mary met John three years before she met Bill.*

Such constructions would lend support to del Prete's comparative analysis, as the construction is reminiscent of degree comparatives as in *Mary is three inches taller than John*.

But how should sentences like (42) be interpreted? I propose that the temporal measure phrase modifies the time t' by moving it along the time axis by the specified distance, as in (43.a). This entails that *before* has a reading that allows for the specification of a temporal distance, as in (43.b), quite similar to local prepositions like *behind*, which allow for distance phrases as in *three meters behind the door*.

- (43) a. $\llbracket \textit{three years [before B]} \rrbracket = \lambda t \neg \exists t' [t' \leq [t+3y] \wedge \llbracket B \rrbracket(t')]$
 b. $\llbracket \textit{before} \rrbracket = \lambda p \lambda d \lambda t \neg \exists t' [t' \leq t+d \wedge p(t')]$

The + operation should work as expected; for example, $1.1.2000+3y = 1.1.2003$. The meaning of $\llbracket A [\textit{three years [before B]}] \rrbracket$ can be represented graphically as follows. As before, A is true at t , and B cannot be true at the dotted line.



We get the following interpretation for (42):

- (45) $\lambda t [t < n \wedge \text{Mary met John at } t \wedge \neg \exists t' [t' \leq [t+3y] \wedge [t' < n \wedge \text{Mary met Bill at } t']]]$

If asserted, this states that there is a time t before the utterance time such that Mary met John at t , and that there is no time t' up to three years after t such that Mary met Bill at t' . This interpretation seems to be not quite sufficient, as it does not entail that Mary actually did meet Bill, and that she met him three years after t . In general, we observe that *before* sentences with measure phrases are factive. For example, (46) is strange, even if Mozart planned to finish his requiem on December 8, 1791, three days after he in fact died, and hence the probable alternative worlds would license the sentence according to Condoravdi and Ogihara.

(46) #*Mozart died three days before he finished his requiem.*

The current proposal can explain the oddness of such sentences, and hence their factivity, in a straightforward way: If the B sentence is not true at all, then any temporal distance degree could be supplied, and the $\llbracket A \textit{ before B} \rrbracket$ sentence would still be true. It also explains why (42) is understood as saying that the time of Mary meeting Bill is precisely three years after the time of Mary meeting John: If it were later, another measure expression denoting a larger amount would have been used, as this would be more informative in this case. Furthermore, it is predicted why sentences with stative verbs, as in *Mary knew John three years before she knew Bill*, are odd: Referring to protracted times, many choices could be made for the measure phrase, which makes any such choice uninformative.

Before clauses with temporal measure phrases license NPIs, as in (47):

(47) *Mary met John three years before she met any of his friends.*

This sentence is true, for example, in case Mary met John in 2000, and met the first one of his friends in 2003. Under the proposed interpretation, expressions like [*3 years before B*] are downward-entailing in B, simply because they are still under the scope of a negation. This predicts that they license NPIs.

The operator *after* also occurs with measure phrases:

(48) *Mary met Bill three years after she met John.*

After with a measure phrase cannot be just the negated form of *before* with the same measure phrase, which would be (49). Rather, we have to assume the meaning specified in (50.a) or equivalently (b), which gives us the representation (51).

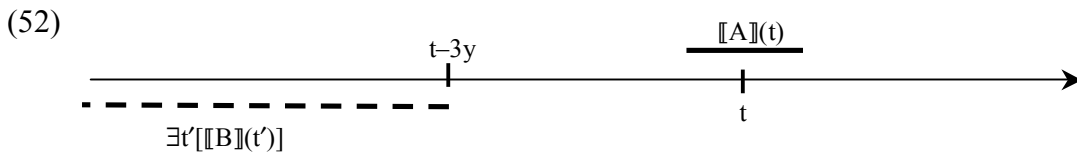
(49) $\llbracket \textit{three years [after B]} \rrbracket = \lambda t \exists t' [t' \leq [t+3y] \wedge \llbracket B \rrbracket (t')]$

(50) a. $\llbracket \textit{three years [after B]} \rrbracket = \lambda t \exists t' [t' \leq [t-3y] \wedge \llbracket B \rrbracket (t')]$

b. $\llbracket \textit{three years [after B]} \rrbracket = \lambda t \exists t' [[t'+3y] \leq t \wedge \llbracket B \rrbracket (t')]$

(51) $\lambda t [t < n \wedge \text{Mary meets Bill at } t \wedge \exists t' [[t'+3y] \leq t \wedge t' < n \wedge \text{Mary meets John at } t']]$

This ensures that if [*A three years before B*] is true, then [*B three years after A*] is true. We can represent the reading of [*A three years after B*] as follows, where A is true at t, and B is true at some time of the hyphenated line.



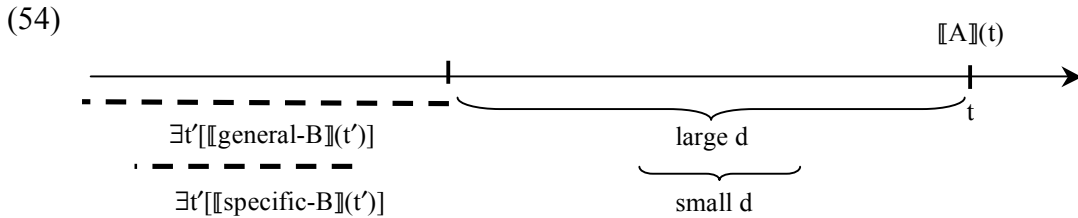
We now turn to the cases in which *after* licenses NPIs, as in (5). Condoravdi's proposal assumes that *after* licenses NPIs after coercion of the *after*-clause with MAX, which entails that the main clause ceased to be true and creates a downward-entailing context. This does not explain, however, why the NPI licensing of *after* requires the presence of temporal measure terms like *long* (cf. Linebarger 1987). There is no reason why temporal measure terms could not apply in case of a coercion with EARLIEST. The present account gives an independent reason why, in the case of [*A d after B*], where d is a temporal measure phrase, the B clause does not hold at the time at which the A clause is evaluated, as in this case the specification of a measure term would be uninformative.

Condoravdi observes that not just any temporal measure phrase helps to license NPIs, which can be appreciated with the following examples:

- (53) a. *Jo kept writing poems long after there was any hope to get them published.*
 b. *Jo kept writing poems many years after there was any hope to get them published.*
 c. *#Jo kept writing poems three years after there was any hope to get them published.*
 d. *##Jo kept writing poems three days after there was any hope to get them published.*

I suggest that this phenomenon comes about as follows. Sentences like (53.a,b) want to make a strong statement. Following the hypotheses concerning the use of NPIs in Krifka (1995), this is achieved by using a proposition as *after*-clause that is semantically weak compared to alternatives,

in the sense that it applies to a broader time span than its alternatives. The choice of a temporal measure phrase indicating a long time period is consonant with this rhetorical goal, and thus these two devices happily go together. For example, (53.a,b) makes the point that Jo’s habit of writing poems went on for a very long time, that is, that A is true at times t that are very late. This is enforced by choosing propositions B that apply to a broader (in particular, later) set of times, and a temporal measure phrase that denotes a longer time.



In principle, speakers can rely only on the first of these devices, a semantically weak proposition; hence (55) is fine as well.

(55) *Jo kept writing poems after there was any hope to get them published.*

On the other hand, as Condoravdi observes, not any NPI is felicitous. It must be an NPI that yields a broader range of times. This is why (56) is bad; as soon as Jo retired to some Caribbean island, then he also retired to “any” island, as the background assumption is that one retires only once.

(56) **Jo wrote poems long after he had retired to any Caribbean island.*

Condoravdi explained the ungrammaticality of (56) by the non-cumulativity of *retire*, which makes it impossible to apply MAX in Condoravdi’s definition. But it appears that there is no other motivation for MAX to be thus restricted. Furthermore, there are other achievement or accomplishment verbs, like *discover*, which have the required temporal property, and which can be used without problem:

(57) *Jo wrote poems long after he had discovered anything new to say.*

In closing this section on *before* and *after* with temporal measure phrases, one could ask whether, in addition to the interpretation that was outlined here, there is another, simpler one that just states that the times at which the A and B clauses are interpreted are temporally separated by the specified time.

(58) a. $[[A \text{ d before } B]] = \lambda t[A(t) \wedge \exists t'[t <_d t' \wedge B(t')]]$

b. $[[A \text{ d after } B]] = \lambda t[A(t) \wedge \exists t'[t >_d t' \wedge B(t')]]$

The factual interpretation of B that we have observed with temporal measure phrases is built into these interpretations directly. Also, *before* and *after* would be converses of each other, as indicated by the $<_d$ and $>_d$ relations, which are supposed to hold between times t, t’ such that t is either d-long before or d-long after t’. The two interpretations appear to be truth-functionally identical, and it certainly could be assumed that they exist side by side. Perhaps real speakers use Occam’s razor, the principle of not assuming different meanings if not necessary, more rarely than linguists do in analyzing language.

4. *Since* and *Until*

The temporal operators *since* and *until* license NPIs only in particular constructions. Condoravdi interprets *until* as a relation between a time and intervals extending up to that time, and *since* as a relation between a time and intervals extending from that time:

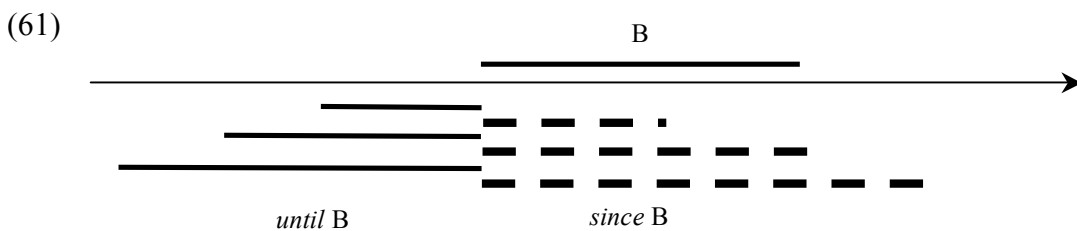
- (59) a. $\llbracket \textit{until} \rrbracket(t')$ = $\lambda t[t < t' \wedge \text{convex}(t \sqcup t')]$
 b. $\llbracket \textit{since} \rrbracket(t)$ = $\lambda t[t' < t \wedge \text{convex}(t \sqcup t')]$

The sum of two times $t \sqcup t'$ is convex if it forms an interval. This entails that there is no gap between t and t' . The time t' is specified by the B sentence in $[A \textit{ until/since} B]$. For *until*, t' could either be the earliest time or the maximal time at which B is true, with no difference in meaning. *John stayed awake until Mary was asleep* means that John stayed awake for a time t that abuts the time at which Mary sleeps for the first time, which is the same as saying that it abuts the whole sleeping time of Mary. Interestingly, for *since*, it is always the earliest time at which B is true. *John stayed awake since Mary had been asleep* means that John stays awake from the time Mary fell asleep onwards, not that he stayed awake for a time after the total time that Mary was asleep. Does this mean that we do need the EARLIEST operator after all?

No, because one can give an alternative interpretation that relates the meanings of *until* and *since* to the meanings of *before* and *after* in the following way: A time qualifies for $[\textit{until} B]$ iff it is before B and there is no later time that is before B; a time qualifies for $[\textit{since} B]$ iff it is after B and there is no earlier time that is after B. Hence *until* and *since* denote maximal before- and after-times:

- (60) a. $\llbracket \textit{until} B \rrbracket$ = $\lambda t[\llbracket \textit{before} B \rrbracket(t) \wedge \forall t'[\llbracket \textit{before} B \rrbracket(t') \rightarrow \neg t < t']]$
 b. $\llbracket \textit{since} B \rrbracket$ = $\lambda t[\llbracket \textit{after} B \rrbracket(t) \wedge \forall t'[\llbracket \textit{after} B \rrbracket(t') \rightarrow \neg t' < t]]$

According to these meanings, $[A \textit{ until} B]$ is true iff A is true at some time before B becomes true and stays true at least up to the point where B becomes true; and $[A \textit{ since} B]$ is true if A is true at least from the point on at which B becomes true. This is achieved without any resort to coercion operators like EARLIEST or MAX.



Both under Condoravdi's assumption and the one developed here, we should not expect NPIs to occur in the scope of *since* or *until*. Take *until*: If Mary sleeps from 23:00 onwards, then *until Mary sleeps* is true for times that end in 23:00; if Mary sleeps deeply from 24:00 onwards, then *until Mary sleeps deeply* is true for times that end in 24:00; one set of these times is not a subset of the other. We have a similar situation with *since*. However, we can find NPIs in the scope of *until*, as the following internet findings suggest:

- (62) a. *Then I waited and waited and waited until anyone got up.*
 b. *Just yesterday I tried hiding in a small Wendy's bag and waited until anyone would walk by and jump out at them then hide in my bag again!*

Such examples show that NPI licensing cannot be reduced to mere downward-entailingness. Downward-entailingness is a sufficient, but not a necessary property for contexts that endorse NPIs. The proper characterisation is that NPIs are licensed in positions where the speech act based on the NPI leads to a better argumentative value than speech acts based on alternatives to the NPI. In (62.a), it is suggested that the speaker had to wait for a long time. The *until* clause expresses a less specific proposition ('a person got up') than its alternatives ('a particular person got up'), where the less specific proposition potentially comes with an earlier starting time; yet the sentence states that the wait was still very long. In (62.b), the underlying expectation seems to be that the speaker jumps out only when particular persons walk by; with respect to that background, to say that one jumps out when any person walks by makes a stronger statement.

Condoravdi discusses another type of case in which *since* and *until* occur with NPIs (cf. also Iatridou 2003):

- (63) a. *It will be five years until Mary will visit any museum.*
 'It will be five years until Mary will visit the first museum.'
 b. *It has been five years since Mary has visited any museum.*
 'It has been five years since Mary visited the last museum.'

Here it is crucial that the measure phrases are downward-bounded; that is, *five years* is to be interpreted as *at least five years*. Then the entailments characteristic for NPI licensing go through:

- (64) *It will be at least five years until Mary will visit the Louvre or the Centre Pompidou*
 ⇒ *It will be at least five years until Mary will visit the Louvre.*
 (65) *It has been at least five years since Mary has visited the Louvre or the Centre Pompidou*
 ⇒ *It has been at least five years since Mary visited the Louvre.*

Again I think we do not need EARLIEST, provided that *until* and *since* are defined as in (60).

5. Conclusion

The major theoretical point of Condoravdi's article is that "quantification over temporal clauses and over main clauses is connective-independent and the presuppositional implication of temporal clauses is to be attributed to the operators operating on temporal clauses prior to their composition with temporal connectives." Condoravdi proposed two such operators, EARLIEST and MAX. I have argued that there is little independent motivation for assuming EARLIEST, and that we should rather not assume coercion in the case of temporal operators like *before*, *after*, *until* and *since*. I also have argued that these operators come with certain implicatures, which help to explain their factual and non-factual use, without the assumption that they have a modal interpretation which would require us to consider other possible worlds – contrary to Ogihara (1995).

There would of course be more to say about the operators discussed here. For example, there is a teleological interpretation of German *bevor* that can be explained quite naturally, based on the non-factual interpretation of *bevor*.

- (66) *Nimm die Hand weg bevor du dich noch verbrennst.*
 'Take off your hand before you burn yourself.'

There are also speechact-related uses that are reminiscent of so-called biscuit conditionals:

- (67) a. *Before you despair, there is a way to address that problem as well, by setting a read/write timeout value via the `stream_set_timeout()` function.*
 b. *But since you asked, there is a type of wire that has another couple of wires warped [sic] around it lengthwise.*

Also, there are factive uses that identify a common ground:

- (68) a. *Since John loves books, you could give him a book as present.*
 b. *Nachdem John Bücher mag, könntest du ihm ein Buch schenken.*
 ‘(Lit.) After John likes books, you could give him a book as present.’

Furthermore, there is a curious use of negation with *before* clauses in which the negative element appears to be “superfluous”, not contributing to the meaning of the overall sentence:

- (69) *Wir reisen nicht ab bevor wir nicht den Louvre gesehen haben.*
 ‘We do not leave before we have [not] seen the Louvre’

See Krifka (2010) for an attempt to explain the occurrence of this negation, making use of the implicature of [A *before* B] that B will become true at some later time.

No doubt it will be quite some time before *before* and *after*, and then some more until *until* and *since*, have revealed all of their secrets.

References

- Anscombe, G.E.M. (1964). “Before and after.” *The Philosophical Review* 74, 3 – 24.
- Beaver, David & Cleo Condoravdi (2003), “A uniform analysis of *before* and *after*”, *SALT* 13, 37-57. Cornell, Ithaca: CLC Publications.
- Chierchia, Gennaro (1998), “Reference to kinds across languages”, *Natural Language Semantics* 6, 339-405.
- Condoravdi, Cleo (this volume), “NPI licensing in temporal clauses”, *Natural Language and Linguistic Theory*.
- del Prete, Fabio (2008), “A non-uniform semantic analysis of the Italian temporal connectives *prima* and *dopo*”, *Natural Language Semantics* 16, 157-203.
- Dowty, David R. (1977), “Toward a semantic analysis of verb aspect and the English ‘imperfective’ progressive”, *Linguistics and Philosophy* 1, 45-77.
- Fernald, Theodore B. (1999), “Evidential coercion: Using individual-level predicates in state-level environments.”, *Studies in the Linguistic Sciences* 29, 43-63.
- Iatridou, Sabine (2003), “A little bit more on the English perfect.” In A. Alexiadou, M. Rathert & A. von Stechow (eds.), *Perfect explorations*, 133-151. Berlin: Mouton de Gruyter.
- Krifka, Manfred (1995), “The semantics and pragmatics of polarity items”. *Linguistic Analysis* 25, 209-257.
- Krifka, Manfred (2010). “How to interpret ‘expletive’ negation under *bevor* in German”, in Th. Hanneforth & G. Fanselow (eds.), *Language and Logos: Festschrift for Peter Staudacher on his 70th Birthday*, Berlin: Akademie Verlag.
- Linebarger, Marica (1987). “Negative polarity and grammatical representation”. *Linguistics and Philosophy* 10, 325-387.
- Moens, Marc & Mark Steedman (1988), “Temporal ontology in natural language.”, *Computational Linguistics* 14, 15-28.

- Ogihara, Toshiyuki (1995). “Non-factual *before* and adverbs of quantification.” *SALT V*, 273-291.
Cornell, Ithaca: CLC Publications.
- Partee, Barbara (1987), “Noun phrase interpretation and type-shifting principles”, in J. Groenendijk,
Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers,
Dordrecht, Foris, 115-143.
- Penka, Doris (2007), *Negative Indefinites*, Ph.D. Dissertation, Universität Tübingen, Tübingen.
- Seuren, Peter, 1973, “The Comparative”, in F. Kiefer & N. Ruwet (eds.), *Generative Grammar in Europe*, Dordrecht: Reidel.
- von Stechow, Kai (1999). “NPI licensing, Strawson entailment”. *Journal of Semantics* 16, 97-148.

Manfred Krifka
Zentrum für Allgemeine Sprachwissenschaft, Berlin
Humboldt-Universität zu Berlin
krifka@rz.hu-berlin.de

Address:
Institut für deutsche Sprache und Linguistik
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin