On the Universal Principles of Tense Embedding: The Lesson from Before

YAELE SHARVIT
University of California, Los Angeles

Abstract

Languages that are classified as non-Sequence-of-Tense come in more than one variety (e.g., Arregui & Kusumoto 1998): some of these languages allow a past tense in before-clauses while others do not. We propose that some languages have quantificational (existential) tenses, while others have pronominal (referential/bound) tenses. The past tense in before-clauses is ill-formed in a language that has quantificational tenses, because the semantics of before is incompatible with existential quantification over times. A language with pronominal tenses does not have this problem. The pronominal/quantificational tense distinction interacts with the Sequence-of-Tense parameter, providing a general theory of possible and impossible languages.

1 INTRODUCTION: TWO KINDS OF NON-SOT LANGUAGES

As shown in (1)–(2), in English before-clauses past-under-past (a past tense in the embedded clause, where the matrix verb is inflected for past tense) is grammatical, and present-under-past (a present tense in the embedded clause, where the matrix verb is inflected for past tense) is ungrammatical. In Japanese (see (3)–(4)), past-under-past is ungrammatical, and present-under-past is grammatical. As shown by (5) and (6), Polish exhibits the English pattern in before-clauses (for a detailed discussion of the data, see Ogihara 1996, Arregui & Kusumoto 1998 and references cited there).

(1) John left the meeting [before Mary arrived].
(2) *John left the meeting [before Mary arrives].
(3) *Taroo-wa [Hanako-ni at-ta mae-ni] denwa-o si-ta
   Taro-TOP Hanako-DAT meet-PAST before phone-ACC do-PAST
   Lit: ‘Taro phoned before (he) met Hanako’
The Polish before-facts are somewhat unexpected, because Polish patterns with English in before-clauses, but in attitude reports it patterns with Japanese. English is traditionally classified as a Sequence-of-Tense (SOT) language, whereas Japanese and Polish are both classified as non-SOT languages. This classification refers to the fact that past-under-past in the English (7) can be understood as “null”, giving rise to the so-called “simultaneous” reading; as opposed to past-under-past in the Japanese (8) and the Polish (9): in Polish and Japanese the “simultaneous” reading must be expressed with an embedded present. The facts in (1)–(9) are summarized in (10).

(7) John said that Mary was sick.
   John: “Mary is sick” ("simultaneous" reading)
   John: “Mary was sick” ("back-shifted" reading)

(8) a. Taroo-wa [Hanako-ga byooki-da] -to it-ta
    Taro-TOP Hanako-NOM be-sick-PRES that say-PAST
    Taro: “Hanako is sick” ("simultaneous" reading)

   b. Taroo-wa [Hanako-ga byooki-datta] -to it-ta
    Taro-TOP Hanako-NOM be-sick-PAST that say-PAST
    Taro: “Hanako was sick” ("back-shifted" reading)

(9) a. Ania powiedziała ze Marcin jest chory
    Ania say-PRF-PAST that Marcin be-PRES sick
    Ania: “Marcin is sick” ("simultaneous" reading)

   b. Ania powiedziała ze Marcin był chory
    Ania say-PRF-PAST that Marcin be-PAST sick
    Ania: “Marcin was sick” ("back-shifted" reading)

(10) | Past-under-past | Pres-under-past | Past-under-past | Pres-under-past |
     | "null" in attitudes | "null" in attitudes | OK in before-clause | OK in before-clause |
     | English | + | -- | + | -- |
     | Japanese | -- | + | -- | + |
     | Polish | -- | + | + | -- |
These facts present a puzzle: one might expect a state of affairs where Japanese and Polish tenses behave in a parallel fashion not just in attitude reports (but also in before-clauses), or a state of affairs where English and Polish behave in a parallel fashion not just in before-clauses (but also in attitude reports).

We attempt to solve this puzzle with the following assumptions: (a) English has an SOT rule (which makes a past-under-past semantically vacuous) while Polish and Japanese do not (von Stechow 1995; Ogihara 1996; Grønn & von Stechow, to appear, among many others); (b) English and Polish have pronominal tenses, roughly in the sense of Partee (1973), while Japanese has quantificational tenses, roughly in the sense of Prior (1967) (cf. Ogihara & Sharvit 2012); and (c) the semantics of before and of say are the same cross-linguistically. The first assumption explains why in English, but not in Japanese or Polish, past-under-past can be “null” in attitude reports. The second and third assumptions “conspire” to explain the cross-linguistic behavior of before.

We assume, with Beaver & Condoravdi (2003), that a before-clause – [before p] in ‘q before p’ – is interpreted as containing a definite description of the form “the first p-time”. We claim that when p contains a quantificational past tense, the definite description cannot pick out an appropriate referent, and this results in a presupposition failure. For example, when (1) has the LF in (11), where the embedded past is an existential quantifier over times, THE-FIRST [PAST [Mary arrive]] means “the first time t such that there is a time t’ preceding t such that Mary arrives at t’”.

(11) [PAST [John leave... before (THE-FIRST [PAST [Mary arrive]])]]

This definite description cannot pick out a unique time even if Mary arrives only once, because of the density of the time axis (that the time axis is dense means that for any two temporal points m1 and m2 such that m1 precedes m2, there is a temporal point m3 such that m1 precedes m3 and m3 precedes m2). For any time t such that Mary’s arrival time precedes t, there is a time t’ preceding t such that Mary’s arrival precedes t’. Given this, no language (that lacks an SOT rule) can have a quantificational past tense in a before-clause.

So why do some languages (that lack an SOT rule) allow past-under-past before-clauses? Borrowing an idea from Ogihara & Sharvit (2012) and Sharvit & Tieu (2011), we propose that while both Japanese and Polish are non-SOT, the former has quantificational tenses while the latter has pronominal tenses, which can be interpreted
either as referring expressions or as bound variables. A pronominal past in a before-clause avoids presupposition failure because it lacks existential quantification. Thus, the definite description that we get with a pronominal past is “the first time t preceding the speech time such that Mary arrives at t”, which picks out the leftmost past arrival time.

It is not commonly thought that quantificational and pronominal tenses give rise to such radically different interpretations. There is a popular idea in the literature (e.g., von Fintel 1994) according to which a quantifier takes a pronominal argument that restricts its domain of quantification. This idea seems to imply that the two treatments of tense – pronominal and quantificational – are (linguistically) completely interchangeable. Yet our claim is that despite similarities between treatments of tenses as pronouns and treatments of tenses as restricted quantifiers, the distinction between (pronominally or non-pronominally) restricted quantifiers and “real” pronouns is significant enough so as to predict the existence of several types of non-SOT languages (as well as several types of SOT languages).

Section 2 provides a brief and informal discussion of the semantics of before. Section 3 shows how the pronominal/quantificational tense distinction interacts with the semantics of before, and Section 4 shows how it interacts with the SOT parameter. Section 5 argues against an alternative theory of before based on Geis-ambiguities (discussed in Geis 1970), and Section 6 proposes a general typology of languages based on the proposal outlined in Sections 3 and 4.

2 BEFORE—AN INFORMAL SEMANTICS

Let us begin by considering some basic properties of past-under-past English before-clauses (see Anscombe 1964; Heinämäki 1974; Stump 1985; Landman 1991; Ogihara 1995, 1996; Beaver & Condoravdi 2003, among others). A past-under-past sentence of the form ‘q before p’ has the properties described in (12), as illustrated by the sentences in (13).

\[
\begin{align*}
\text{(12) a. } & \text{q-factivity: } \text{Some time } t \text{ is a } q\text{-time.} \\
& \text{b. } q\text{-precedence: } \text{Some } q\text{-time precedes any and all } p\text{-times.} \\
& \text{c. NPI-licensing in } p: \text{The } p\text{-clause may contain a negative polarity item (NPI).}
\end{align*}
\]

\[
\begin{align*}
\text{(13) a. } & \text{John watered the plant before it died.} \\
& \text{b. } \text{John watered the plant before it ever got sick.}
\end{align*}
\]

Suppose SP = the speech time, q = \{t| John waters the plant at t and t is prior to SP\}, and p = \{t| the plant dies at t and t is prior to SP\}.
Intuitively, (13a) entails that: (i) John waters the plant prior to SP ((12a)), and (ii) the watering time precedes any and all dying times ((12b)). The acceptability of (13b) – with $\textit{ever}$ in $p$ – illustrates (12c) (cf. *John ever watered the plant before it got sick).

It is helpful to test existing theories of before against (12a–c), such as Ogihara (1996) and Beaver & Condoravdi (2003), which differ from each other in their treatments of $p$, as shown in the informal (14) and (15).

(14) Ogihara-style ‘before’: ‘$q$ before $p$’ is true iff some $q$-time precedes some $p$-time.  
(15) B&C-style ‘before’: ‘$q$ before $p$’ is true iff some $q$-time precedes the first $p$-time.

An Ogihara-style analysis introduces a $p$-time “existentially”; a B&C-style analysis refers to a (presumably presupposed) first $p$-time. Both (14) and (15) account for $q$-factivity (and, to some extent, both account for $q$-precedence). But (15) has a clear advantage over (14) regarding NPI-licensing, because it implies that before is downward entailing on $p$ in the sense of Ladusaw (1979) (or, more accurately, Strawson downward entailing on $p$ in the sense of von Fintel 1999), correctly predicting NPIs to be licensed in the before-clause (see Beaver & Condoravdi 2003, Condoravdi 2010): for any $p$ and $p'$ (e.g., “I went to Buenos Aires” and “I went to Buenos Aires to have a good time” respectively) such that $p'$ entails $p$, any $t$ that precedes the first $p$-time also precedes the first $p'$-time (assuming there is a first $p'$-time). (14) implies that before is not downward entailing on $p$.\footnote{Anscombe’s (1964) analysis (see also Landman 1991; and Arregui & Kusumoto 1998), with universal quantification over $p$-times (‘$q$ before $p$’ is true iff there is a $q$-time $t$ and for all $p$-times $t'$, $t$ precedes $t'$), also renders before downward entailing. However, Beaver & Condoravdi (2003) point out that $\textit{John left exactly two minutes before Sally arrived}$, where the before-clause is preceded by a degree phrase (see Appendix), is incompatible with universal quantification over $p$-times Sally can arrive more than once, but John’s leaving precedes only her first arrival by exactly two minutes.} For us, this is a deciding factor, and we opt for (15).

Notice that the times that ‘$q$ before $p$’ talks about are contextually restricted, as implied by examples such as (Today,) John didn’t sing before Mary danced. This sentence is judged true when the context is such that only “today” is relevant (even when $\textit{today}$ is not pronounced), as long as today there is only one singing event by John and one dancing event by Mary, the latter precedes the former, and yesterday Mary didn’t dance but John sang. To account for this contextual restriction, we replace (15) with (16), where $C$ is a set of contextually relevant times.

(16) ‘$q$ before $p$’ is true iff some $q$-time in $C$ precedes the first $p$-time in $C$. 

\footnote{Anscombe’s (1964) analysis (see also Landman 1991; and Arregui & Kusumoto 1998), with universal quantification over $p$-times (‘$q$ before $p$’ is true iff there is a $q$-time $t$ and for all $p$-times $t'$, $t$ precedes $t'$), also renders before downward entailing. However, Beaver & Condoravdi (2003) point out that $\textit{John left exactly two minutes before Sally arrived}$, where the before-clause is preceded by a degree phrase (see Appendix), is incompatible with universal quantification over $p$-times Sally can arrive more than once, but John’s leaving precedes only her first arrival by exactly two minutes.}
Restricting the relevant times to C allows the exclusion of any irrelevant q-times (John’s singing yesterday) that exist prior to the relevant p-time (Mary’s dancing today). So in our example, a plausible value for C is such that $C \subseteq \{t| t \text{ is contained in “today”}\}$.

Notice also that (13a) may have a factive reading (according to which the plant dies prior to SP) or a non-factive reading (according to which the plant does not die prior to SP, or at all). We call this property of English past-under-past before-clauses ‘No-p-factivity’. The oddity of examples such as (17) – as opposed to the felicity of (13a) when the plant doesn’t die – suggests that No-p-factivity comes from the intensionality of before (on this, see Landman 1991; Ogihara 1995; Beaver & Condoravdi 2003; Condoravdi 2010).

(17) #John watered the plant before a unicorn destroyed it.

So we replace (16) with (18).

(18) ‘q before p’ is true in world w iff there is a t in C such that t is a q-time in w and for all worlds w’ accessible from w at t, t precedes the first element of $\{t| t \text{ is in C and t is a p-time in w’}\}$.

Assuming that the worlds accessible from w are similar to w (at least regarding whether they contain unicorns), (17) is predicted by (18) to be odd when evaluated relative to any world w that doesn’t have unicorns in it. Furthermore, (18) accounts for No-p-factivity in general, because the truth of (13a) now does not require that the plant die in w.

Wherever possible, and mainly for simplicity, we use an “extensional” semantics for before – i.e., some variant of (16) – that delivers the factive reading. In the next section we propose a more formal version of (16), and account for the cross-linguistic variation presented in Section 1.

3 THE PROPOSAL

In 3.1 we show that if we take tenses to be existential quantifiers and interpret both occurrences of the past tense in (13a) – matrix and embedded – using a B&C-semantics for before, we end up with no denotation for the before-clause (see also Sharvit & Tieu 2011). In 3.2 and 3.3 we show that the problem disappears if the embedded past is pronominal; thus, the English/Japanese contrast regarding before-clauses (see (10)) can be explained if we assume that the English past tense is pronominal and the Japanese past tense quantificational.
3.1 The quantificational past problem

We start by showing that a quantificational past tense in a before-clause results in a fatal presupposition failure. The idea is simple: assuming that the past is an existential quantifier over times, and that before shifts the evaluation time of that quantifier, we derive for John watered the plant before it died the following meaning: “the watering of the plant by John precedes the first time t such that there is a time t’ that precedes t such that the plant dies at t’”. On the assumption that the time axis is dense, for any time t such that the dying precedes t, there is a time t’ preceding t such that the dying precedes t’. Consequently, there is no “first” time that meets the relevant description (the inspiration for this idea comes from Fox & Hackl’s 2006 treatment of scalar implicatures). Since this is an essential part of our proposal (specifically, we take it to be the reason why past-under-past is not possible in Japanese before-clauses), we now discuss it in some detail.

As in the tradition established by Bennett & Partee (1972), we distinguish between M – the set of all moments, which is homomorphic to the interval (–∞, +∞) (the set of real numbers) and is therefore itself an interval – and D_i, which is the set of all closed sub-intervals of M, i.e., the set of all sub-intervals of M of the form [m1, m2], where m1 and m2 are elements of M and m1 ≤ m2 (‘<’ stands for ‘precedes’ and ‘≤’ stands for ‘precedes or equals’). Elements of D_i serve as arguments of verbal predicates (the classes these predicates belong to determine what kind of intervals they may take; often, maybe always, non-singletons). The statement in (19) follows from the homomorphism between M and (–∞, +∞) and the definition of D_i.

(19) For every t ∈ ({M} ∪ D_i):
(i) if t is a non-singleton, t is an infinite set of moments;
(ii) the elements of t are ordered by ‘≤’ (i.e., for any m1,m2 ∈ t such that m1 ≠ m2: m1 < m2 or m2 < m1); and
(iii) t is dense (i.e., for any m1,m2 ∈ t such that m1 < m2: there is a m3 ∈ t such that m1 < m3 and m3 < m2).

Let us, temporarily, assume a more-or-less standard quantificational analysis of English tenses, according to which the past tense is a sentential operator that expresses existential quantification over times. In this system, PAST^{ENG} is of semantic type <<i, t>, <i, t>>; in other words, the English past tense denotes an element of D_{<<i,t>,<i,t>>>} (D_i = {True, False}). The sentential argument of PAST^{ENG} is p in (20); its temporal argument – t in (20) – is its evaluation time interval. g is a variable assignment, and ‘<’ is generalized to pairs of elements of
Di, and pairs consisting of an element of Di and an element of M; see (21).

(20) For any t ∈ Di and p ∈ D, [[PAST\_ENG]]\(p(t)\) is defined only if there is a t’ ∈ Di such that t’ < t and p(t’) is defined.
When defined, [[PAST\_ENG]]\(p(t)\) := True iff there is a t’ ∈ \{t'' < t | p(t'') is defined\} such that p(t’) = True.2

(21) a. For any t, t’ ∈ Di: t < t’ iff MAX(t) < MIN(t’);
for any t ∈ Di and m ∈ M: m < t iff m < MIN(t), and t < m iff MAX(t) < m.
b. For any S ⊆ M:
\(\text{MIN}(S)\) is the unique m ∈ S such that for all m’ ∈ S, m ≤ m’, if there is such a unique m (otherwise, \(\text{MIN}(S)\) is undefined); and
\(\text{MAX}(S)\) is the unique m ∈ S such that for all m’ ∈ S, m ≥ m’, if there is such a unique m (otherwise, \(\text{MAX}(S)\) is undefined).

In addition we assume, as is generally assumed, that [[John]]\(g\) and [[the plant]]\(g\) are both elements of De, the domain of individuals, and [[water]]\(g\) is an element of D<e,<<e,<<i,t>>>. Accordingly, the LF of John watered the plant is (22a), where John water the plant – the internal argument of PAST\_ENG – is of type <i, t> (water combines with the plant and John by Functional Application, as stated, for example, in Heim & Kratzer 1998), and t0 – the external argument of PAST\_ENG – is a silent pronoun denoting an element of Di (cf. Kusumoto 1999).3 By convention, whenever g is the assignment supplied by the speech context, g(0) = SP. PAST\_ENG combines with John water the plant and with t0 by Functional Application, yielding the meaning in (22b), repeated in a simplified format in (22c), where the presupposition ‘t < SP’ restricts the domain of quantification to past time intervals.

(22) a. \[t_0 [PAST\_ENG [John water the plant]]]\n
b. Whenever defined, [[22a]]\(\text{SP}^\text{SP}\) = True iff there is a t ∈ \{t’ < SP | [[John water the plant]]\(p(t)\) is defined\} such that [[John water the plant]]\(p(t)\) = True.
c. There is a t < SP such that John waters the plant at t.

2 We assume that an indefinite is restricted by the presuppositions in its scope (cf. Heim 1998b), as implied by (i). Since PAST is assumed to be an indefinite, A and B in (i) correspond, respectively, to “t < t’” and p in (20).

(i) [[some A is B]] is defined only if there is a x such that [[A]](x) = True and [[B]](x) is defined. When defined, [[some A is B]] = True iff there is a x ∈ \{y | [[A]](y) = True and [[B]](y) is defined\} such that [[B]](x) = True.

3 Alternatively, PAST\_ENG is of type <i, <i, t>>, t>> and the LF is [PAST\_ENG\_t0 [John [water the plant]]].
Finally, we adopt a more formal rendition of (16), namely, (24), according to which \( \text{before} \) heads a temporal adjunct, and is of type \( <<i, t>, <i, t>> \). (24) refers to \textsc{earliest}, defined in (23): from the contextually supplied time interval \( C \) (recall the discussion of \textit{John didn’t sing before Mary danced}, in connection with (16)), \textsc{earliest} picks the leftmost moment belonging to an element of the set of intervals denoted by the complement of \( \text{before}^{B\&C} \) (cf. von Stechow 2009). According to (24), \( \text{before}^{B\&C} \) presupposes that there is such a leftmost moment, and that it follows the leftmost moment of \( C \) (see Appendix for an intensional alternative to (24), along the lines of (18)).

(23) For any \( t \in D_i \) and \( P \subseteq D_i \), let \( P^t = \{ t' \subseteq t | t' \in P \} \). Then: 
\[ \text{earliest}_t(P) := \text{min}(\cup P^t) \] (see (21b)).

(24) For any contextually-supplied \( C \in D_i \), any \( t \in D_i \) and any \( p \in D_{<i,t>} \), \( [[\text{before}^{B\&C}]^{C,g}(p)(t) \) is defined only if:
(i) \( t \subseteq C \);
(ii) \( \text{earliest}_C(\{ t' \in \text{Dom}(p) | p(t') = \text{True} \}) \) is defined; and
(iii) \( \text{min}(C) < \text{earliest}_C(\{ t' \in \text{Dom}(p) | p(t') = \text{True} \}) \).

When defined, \( [[\text{before}^{B\&C}]^{C,g}(p)(t) := \text{True iff } t < \text{earliest}_C(\{ t' \in \text{Dom}(p) | p(t') = \text{True} \}). \)

(When the presuppositions of \( p \) are significant, we specify what \( \text{Dom}(p) \) is; otherwise, we use ‘\{ t | p(t) = \text{True} \}’ as shorthand for ‘\{ t \in \text{Dom}(p) | p(t) = \text{True} \}’.)

Accordingly, (25) is the LF of (13a): The matrix \( \text{PAST}^{\text{ENG}} \) has widest scope; the embedded \( \text{PAST}^{\text{ENG}} \) takes scope over \textit{it die}.

(25) \[ t_0 [\text{PAST}^{\text{ENG}} [\text{John water the plant} \ [\text{before}^{B\&C} \ [\text{PAST}^{\text{ENG}} \ [\text{it die}]]]]]] \]
\( [[\text{before}^{B\&C}]^{C,g} \) applies to \( [[\text{PAST}^{\text{ENG}} \ \text{it die}]^{C,g} \ - \ an \ element \ of \ D_{<i,t>} \). \textit{John water the plant} and \( \text{before}^{B\&C} \ \text{PAST}^{\text{ENG}} \ \text{it die} \) are both of type \( <i, t> \), and combine by Predicate Modification (as stated, for example, in Heim & Kratzer 1998), with the result that \( [[\text{John water the plant} \ [\text{before}^{B\&C} \ \text{PAST}^{\text{ENG}} \ \text{it die}]]^{C,g} \) is a function \( f \in D_{<i,t>} \) such that every \( t \) in the domain of \( f \) (a domain restricted at least by the definedness conditions of \( \text{before}^{B\&C} \)) is mapped to True just in case \( [[\text{John water the plant}]^{C,g}(t) = [[\text{before}^{B\&C} \ \text{PAST}^{\text{ENG}} \ \text{it die}]^{C,g}(t) = \text{True} \). But (25) does not yield the meaning we are after: we want (25) to yield a meaning along the lines of (26a) (that is to say, we want \textsc{earliest} to pick out the leftmost dying moment within \( C \)); but instead, because of the embedded \( \text{PAST}^{\text{ENG}} \), the meaning actually obtained is (26b).
(26) a. There is a \( t < \text{SP} \) such that:
   (i) John waters the plant at \( t \); and (ii) \( t < \text{earliest}_C(\{t' \mid \text{the plant dies at } t'\}) \)
   (where \( t \subseteq C \) and \( \text{MIN}(C) < \text{earliest}_C(\{t' \mid \text{the plant dies at } t'\}) \)).

b. There is a \( t < \text{SP} \) such that:
   (i) John waters the plant at \( t \); and (ii) \( t < \text{earliest}_C(\{t' \mid \text{there is a } t'' < t' \text{ such that the plant dies at } t''\}) \)
   (where \( t \subseteq C \) and \( \text{MIN}(C) < \text{earliest}_C(\{t' \mid \text{there is a } t'' < t' \text{ such that the plant dies at } t''\}) \)).

The problem with (26b) is that for any \( C \), \( \text{earliest}_C(\{t' \mid \text{there is a } t'' < t' \text{ such that the plant dies at } t''\}) \) is either undefined or equals \( \text{MIN}(C) \); both result in a presupposition failure. We show this in (27), where we consider two cases that cover all the logical possibilities: Case I, where \( \{t < C \mid \text{the plant dies at } t\} = \emptyset \) (i.e., the plant doesn’t die at all or it dies but not prior to \( C \)); and Case II, where \( \{t < C \mid \text{the plant dies at } t\} \neq \emptyset \) (i.e., the plant dies prior to \( C \)). We call this problem Inherent Presupposition Failure (IPF).

(27) Let \( P = \{t' \mid \text{there is a } t'' < t' \text{ such that the plant dies at } t''\} \), and let \( P^C = \{t \subseteq C \mid t \in P\} \).

a. Case I. Since \( C \in D_n \), it follows from (19) that \( C \) is dense. From this it follows that whenever \( \{t < C \mid \text{the plant dies at } t\} = \emptyset \), for every \( t \in P^C \) there is a \( t' \in P^C \) such that \( t' < t \). From this it follows that \( \text{earliest}_C(P) \) is undefined, in violation of presupposition (ii) in (24). (Clearly, it would also follow from (19) that \( \text{earliest}_C(P) \) is undefined if we modified the semantics of \( \text{before}_{B&C} \) so as to allow \( C \) to be any element of \( \{\ associates \cup D_n\} \).

b. Case II. Whenever \( \{t < C \mid \text{the plant dies at } t\} \neq \emptyset \), \( C \in P^C \). From this it follows that whenever \( \{t < C \mid \text{the plant dies at } t\} \neq \emptyset \), \( \text{MIN}(C) = \text{earliest}_C(P) \), in violation of presupposition (iii) in (24).

Case I in (27) represents a problem that comes from the density of \( C \), and one might wonder whether this problem could be avoided by treating \( C \) as a set of (possibly non-overlapping) intervals rather than an interval (in effect, by “digitizing” \( C \), making it non-dense). But this move wouldn’t work: suppose \( \{t1, t2\} \subseteq C \), \( t1 < t2 \), \( t1 \) is the (first, and probably only) dying interval, and there is no other \( t \in C \) such that \( t1 < t \). Then \( \text{MIN}(t2) = \text{MIN}(\cup \{t \in C \mid t \in P\}) \) (where \( P \) is as defined in (27)). The presuppositions of \( \text{before}_{B&C} \) are easily satisfied, but q-precedence (see (12)) is not guaranteed: the watering is required to precede \( t2 \); it is not required to precede the dying – an unwelcome prediction (this problem was noticed in Kusumoto 1999 and Grønn &
von Stechow to appear). Maintaining the assumption that C is an interval easily gets rid of (25) due to IPF.

Case II in (27) represents a problem that comes from presupposition (iii) in (24), and one might wonder whether this problem could be avoided by getting rid of that presupposition. Such a move wouldn’t work either, because of a potential contradiction: t, the time argument of before\textsuperscript{h\&c}, is required by (24) to be contained in C and to precede E\textsubscript{ARIO}ST\textsubscript{C}(P). When E\textsubscript{ARIO}ST\textsubscript{C}(P) = MIN(C), t is required to be both contained in C and precede C – an impossible state of affairs. Moreover, if we also got rid of presupposition (i), which requires t to be contained in C, q-precedence would again not be guaranteed: when E\textsubscript{ARIO}ST\textsubscript{C}(P) = MIN(C), it is required that some watering time precede that moment, but it is not required that any watering time precede the dying time itself.

We therefore leave (24) intact, and assume that (24) is the meaning of before in all languages, predicting that any (25)-like LF leads to IPF (but see Section 5 for discussion of yet another potential way to avoid IPF, namely, the relative-clause analysis of before). This leaves us with the following questions: (a) Why does English allow past-under-past before-clauses? and (b) Why do English and Japanese contrast with respect to past-under-past and present-under-past before-clauses, as we saw in Section 1, and as the following counterparts of John watered the plant before it died/*dies confirm?

(28) Jon-wa [hana-ga kare-u/*ta mae-ni] mizu-o yat-ta
       John-TOP flower-NOM wither-PRES/PAST before-at water-ACC give-PAST

We propose that: (i) Japanese past tense has a semantics along the lines of (20), which leads to IPF in before-clauses; (ii) English avoids IPF in before-clauses by utilizing a pronominal past tense; and (iii) Japanese avoids IPF by utilizing a “null” quantificational present tense in before-clauses.

### 3.2 Pronominal vs. quantificational tenses

As many readers probably know, Partee (1973) showed that an “extreme” quantificational approach to tense (which was the popular approach at the time; see Prior 1967 and Montague 1974) may be problematic, in view of examples such as I didn’t turn off the stove, which seem to make a statement about a specific time (or set of times). On an “extreme” quantificational approach to tense, according
to which quantifiers are either unrestricted or minimally restricted, this example receives either the $\neg \exists$ reading in (29a) or the $\exists \neg$ reading in (29b). Neither reading represents the intuitive meaning of the sentence: the former implies “Speaker has never turned off the stove” (probably false); the latter implies “On at least one past occasion, Speaker didn’t turn off the stove” (almost trivially true).

(29) a. There is no $t < SP$ such that Speaker turns off the stove at $t$.
   b. There is a $t < SP$ such that Speaker doesn’t turn off the stove at $t$.

Instead, Partee suggested that tenses are pronouns that may refer to contextually salient times (and, like more “traditional” pronouns such as he and she, may be bound when in the scope of an appropriate operator). However, as suggested in Bäurle (1977), Partee (1984) and Oghihara (1996), it is possible to address Partee’s concern by treating tenses as contextually restricted quantifiers. Thus, we may treat tenses as pronominal or quantificational (for some concrete suggestions, see von Stechow 1995, 2009; Kratzer 1998; Gronn & von Stechow, to appear, among many others). Given these two theoretical options, we may assume that: (a) a pronominal past tense denotes an element of $D_i$, as defined in (30a), its first index supplies the evaluation time, its second index is its referential index, which supplies its denotation; (b) a restricted quantificational past tense denotes an element of $D_{<i, t>, <i, t>}$, as defined as in (30b) (where $K$ is a contextually supplied time interval). (31a,b) illustrate how the intuitive meaning of I didn’t turn off the stove is obtained, in both approaches.

(30) a. $[[\text{past}_{j,k}]]^g$ is defined only if $g(k), g(j) \in D_i$ and $g(k) < g(j)$.
   When defined, $[[\text{past}_{j,k}]]^g := g(k)$.
   b. For any $K, t \in D_i$, $[[\text{PAST}]]^K g(p)(t)$ is defined only if $K < t$ and there is a $t' \in D_i$ such that $t' \subseteq K$ and $p(t')$ is defined. When defined, $[[\text{PAST}]]^K g(p)(t) := \text{True}$ if there is a $t' \in \{t'' \subseteq K | p(t'') \text{ is defined}\}$ such that $p(t') = \text{True}$.

(31) a. $[\not [\text{past}_{0.3} [I \text{ turn-off the stove}]]$
   Speaker doesn’t turn off the stove at $g(3)$ (where $g(3) < SP$).
   b. $[\not [t_0 [\text{PAST} [I \text{ turn-off the stove}]]]
   There is no $t \subseteq K$ such that Speaker turns off the stove at $t$ (where $K < SP$).

When free, and by convention, the evaluation index of a pronominal tense is interpreted as SP, but in principle both indices could be bound when in the scope of a temporal shifter. Both (31a,b) are viable LFs and both successfully address Partee’s concern. Within more recent work on quantifiers (von Fintel 1994; Martí 2006, among others), which treats the restrictor as a silent pronominal expression rather than an implicit
restrictor, \(PAST\) could be viewed as taking a pronominal restrictor argument, rather than being implicitly restricted (as is the case in (30b) and (31b)).

While it is true that the pronominal/quantificational distinction is not all that significant when it comes to addressing Partee’s concern, it is very significant when it comes to IPF: the absence of an existential quantifier in the scope of before\(_{n\&c}\), as we will soon see, avoids IPF in past-under-past before-clauses. Exploiting this fact, and borrowing an idea from Ogihara & Sharvit (2012) (see also Sharvit & Tieu 2011), we propose that languages may differ from each other with respect to whether their tenses are pronominal or quantificational. More specifically, we propose that Japanese has quantificational tenses and English has pronominal tenses.\(^4\)

The idea that tenses can be pronominal suggests that tenses need not be base-generated in a pre-sentential position (contra what we have assumed), but may instead be internal arguments of verbs. We assume from now on that tenses are indeed direct argument of verbs, though the empirical motivation for this assumption will be made clear only in Section 4. More concretely, we assume that the temporal argument of a verb is its first argument (rather than its last argument). For example, \textit{water} is of type \(<i, e, e, t>\) (rather than \(<e, e, i, t>\)). Consequently, if the syntactic argument of a verb is a quantificational tense (\(PAST\) or \(PRES\), of type \(<<i, t>, <i, t>>\)), it must undergo Quantifier Raising (QR), leaving behind a trace \(-t-\) of type \(i\), to avoid a type-mismatch. QR results in the creation of a \(\lambda\)-abstract (of type \(<i, t>\)). If the syntactic argument of a verb is a pronominal tense (\(past\) and \(pres\), of type \(i\)), it need not undergo QR (later, in Section 4, we will claim that a pronominal tense \textit{cannot} undergo QR; for now, the weaker \textit{need not} suffices). Syntactic \(\exists\)-closure (which, like QR, creates a \(\lambda\)-abstract of type \(<i, t>\)) may optionally apply to a clause with a pronominal tense. Adopting an idea from Lebeaux 1988 and Fox 1999, we assume that adverbial clauses (e.g., before-clauses, which are of type \(<i, t>\)) may be added to an LF at a relatively late stage, crucially after QR or the optional syntactic \(\exists\)-closure have taken place, as illustrated schematically in (32).\(^5\)

\(^4\) Ogihara & Sharvit’s (2012) proposal is slightly different from the current proposal, as it does not rule out the possibility that Japanese has pronominal tenses. But that proposal does not aim to account for before-clauses.

\(^5\) \(\exists\) is interpreted as follows (cf. Fn. 2): \([\exists \alpha]\) is defined only if there is a \(y\) such that \([\alpha](y)\) is defined; when defined, \([\exists \alpha] = True\) iff there is a \(x \in \{y|\alpha(y)\text{ is defined}\}\) such that \([\alpha](x) = True\).

Importantly, while \(\exists\) is itself a quantifier, \(past/pres\) are not, even when bound by \(\exists\): they are elements of \(D_i\).
In addition, we adopt the fairly standard Abstraction Rule in (33) (as stated, for example, in Heim & Kratzer 1998), according to which $[\lambda_n \ldots vbl_n \ldots]$ denotes a (possibly partial) function whose domain is restricted by the presuppositions of $\ldots vbl_n \ldots$.

$$(33) \; [\lambda_n \alpha][x] := [\lambda x : [\alpha][x] \downarrow x] \; \text{(i.e., that function } f \; \text{such that: (i) } \text{Dom}(f) = \{x \mid [\alpha][x] \downarrow x \} \text{ is defined}; \; \text{and (ii) for any } x \in \text{Dom}(f), f(x) = [\alpha][x] \downarrow x).$$

On the assumption that Japanese tenses are quantificational and English tenses pronominal, we generate the LFs in (34) and (35) for the Japanese and English past-under-past counterparts of John watered the plant before it died. (34a) corresponds to (32a), with $[\text{before}_{B&C} PAST_{JAP} [\lambda_2 [it \; die-t_2]]]$ as AdvP, and $[\lambda_1 [John \; water-t_1 \; the \; plant]]$ as its type-identical sister ($PAST_{JAP}$ is an existential quantifier, defined as in (30b); for simplicity the contribution of the contextual restrictor $K$ is suppressed). (35a) corresponds to (32b), with $[\text{before}_{B&C} [\lambda_2 [it \; die-past_{ENG} 0,2]]]$ as AdvP, and $[\lambda_3 [John \; water-past_{ENG} 0,3 \; the \; plant]]$ as its type-identical sister ($past_{ENG}$ is a pronoun, defined as in (30a)).

$$(34) \; a. \; [t_0 \; [PAST_{JAP} [\lambda_1 [John \; water-t_1 \; the \; plant]] \; [before_{B&C} \; PAST_{JAP} [\lambda_2 [it \; die-t_2]]]]]$$

b. There is a $t < SP$ such that: (i) John waters the plant at $t$; and (ii) $t < \text{EARLIEST}_C(t' \mid \text{there is a } t'' < t' \text{ such that the plant dies at } t'')$ (where $t \subseteq C$ and $\text{MIN}(C) < \text{EARLIEST}_C(t' \mid \text{there is a } t'' < t' \text{ such that the plant dies at } t''))$.

$$(35) \; a. \; \exists \; [\lambda_3 [John \; water-past_{ENG} 0,3 \; the \; plant]] [before_{B&C} [\lambda_2 [it \; die-past_{ENG} 0,2]]]$$

b. There is a $t < SP$ such that: (i) John waters the plant at $t$; and (ii) $t < \text{EARLIEST}_C(t' < SP \text{ the plant dies at } t')$ (where $t \subseteq C$ and $\text{MIN}(C) < \text{EARLIEST}_C(t' < SP \text{ the plant dies at } t')$).
Due to the presence of $PAST^{AP}$ in the scope of $before^{B&C}$, (34a) has the same IPF problem as (25).\textsuperscript{6} By contrast, IPF is avoided in (35), because there is no existential quantifier in the scope of $before$: $EARLIEST_C(\{t' < SP\mid \text{the plant dies at } t'\})$ is the leftmost past dying moment within $C$. The presupposition contributed by the embedded past – $past_{0,2}$, whose referential index is abstracted over to create an $<i, t>$-complement for $before^{B&C}$ – restricts the domain of the function denoted by $[\lambda_2 [\text{it die-past}^{ENG}_{0,2}]]$ to intervals that precede $SP$. This follows from the Abstraction Rule in (33). It is expected, then, that any language that has an English-like pronominal past tense allows past-under-past $before$-clauses. It is also expected that any language that has a Japanese-like quantificational tense does not allow past-under-past $before$-clauses.

The LFs of the present-under-past counterparts of the Japanese and English $John$ watered the plant before it dies are as in (36) and (37) respectively, $PRES^{AP}$ has the “null” meaning in (38) (i.e., it makes no significant semantic contribution) and $pres^{ENG}$ has the meaning in (39).

(36) a. $[t_0 [PAST^{AP}[\lambda_1 [John \ \text{water-t}_1 \ \text{the plant}]][before^{B&C} [PRES^{AP} [\lambda_2 [\text{it die-t}_2]]]]]]$

b. There is a $t < SP$ such that: (i) $John$ waters the plant at $t$; and (ii) $t < EARLIEST_C(\{t' \mid \text{the plant dies at } t'\})$ (where $t \subseteq C$ and $MIN(C) < EARLIEST_C(\{t' \mid \text{the plant dies at } t'\})$).

(37) a. $\exists [\lambda_2 [\text{John water-past}^{ENG}_{0,2} \ \text{the plant}]] [before^{B&C} [\lambda_3 [\text{it die-pres}^{ENG}_{0,3}]]]$

b. There is a $t < SP$ such that: (i) $John$ waters the plant at $t$; and (ii) $t < EARLIEST_C(\{t' = SP\mid \text{the plant dies at } t'\})$ (where $t \subseteq C$ and $MIN(C) < EARLIEST_C(\{t' = SP\mid \text{the plant dies at } t'\})$).

\textsuperscript{6} The readers can verify that IPF does not go away in (34) even when the contribution of $K$ is not suppressed. In addition, IPF persists also if we take tense restrictors to be pronominal, as shown in (i) and (ii) (iii) is the “pronominal restriction” variant of (30b).

(i) $[t_0 [PAST^{AP}[\lambda_1 [John \ \text{water-t}_1 \ \text{the plant}]][before^{B&C} [PAST^{AP}-K_3 [\lambda_4 [\text{it die-t}_4]]]]]$

(ii) $[t_0 [PAST^{AP}-K_3 [\lambda_1 [John \ \text{water-t}_1 \ \text{the plant}]][before^{B&C} [\lambda_3 [t_0 [PAST^{AP}-K_3 [\lambda_4 [\text{it die-t}_4]]]]]]]$

(iii) For any $K,t \in D_0$, $[PAST^{AP}]^{C,[K]}(p)(t)$ is defined only if: (i) $K < t$, and (ii) there is a $t' \in D_i$ such that $t' \subseteq K$ and $p(t')$ is defined. When defined, $[PAST^{AP}]^{C,[K]}(p)(t) := True$ iff there is a $t' \in \{t' \subseteq K \mid p(t')$ is defined$\}$ such that $p(t') = True$.

(i) is a case of IPF for the same reason (34a) is. In (ii), the embedded pronominal restrictor is bound by $before$. Whenever (ii) is defined, $EARLIEST_C(\{t \mid [t_0 [PAST^{AP}-K_3 [\lambda_4 [\text{it die-t}_4]]]]^{\text{min}}_{k<3 \ast t;} SP = True\}) = EARLIEST_C(\{t < SP\mid \text{the plant dies at } t'\})$ equals $MIN(C)$. If we remove the “offending” presuppositions from the meaning of $before$, or if we “digitize” $C$ (see discussion of our attempt to rescue (25) in 3.1), $EARLIEST_C(\{t < SP\mid \text{the plant dies at } t'\})$ need not be the leftmost dying time; it could precede it. This matters especially when the $before$-clause contains a degree-phrase, as in $John$ watered the plant exactly two days before it died: we predict that the watering can be separated from $EARLIEST_C(\{t < SP\mid \text{there is a } t' \subseteq t \mid \text{such that the plant dies at } t'\})$ by exactly two days, which implies, counter-intuitively, that the watering can be separated from the dying itself by more than two days (see Appendix for discussion of degree phrases in $before$-clauses).
Because of the semantics of $\text{PRES/pres}$ (both of which lack existential quantification), neither (36) nor (37) gives rise to IPF, and $\text{earliest}$ picks out the leftmost dying time (if there is one). However, in (37) the evaluation index of $\text{pres}$ is ‘0’, which is not bound, so the dying time must coincide with SP. This leads to a pragmatic oddity: the English $\text{John watered the plant before it dies}$ asserts that John’s past watering precedes SP, an assertion that follows from the meaning of the tenses themselves, even without the semantic contribution of $\text{before}^{\text{B&C}}$. Following $\text{Stump}$ (1985), we attribute the unacceptability of the English sentence to this effect, which we call the Stump effect.\(^7\) The Japanese (36) does not cause the Stump effect, as the argument of $\text{before}^{\text{B&C}}$ is the $t_0$-less $[\text{PRES}^{\text{JAP}}_t [\ldots]]$. It is expected, then, that any language with a Japanese-like present allows present-under-past $\text{before}$-clauses, whereas any language with an English-like present disallows such $\text{before}$-clauses (see Section 1).\(^8\)

It is crucial that we make sure that no unwanted pronominal LFs are generated, as (37a) is not the only conceivable pronominal English LF. This calls for some discussion of pronominal tense indexing and binding. First, if $\text{9-closure}$ is completely free, (40), where the evaluation index of the embedded $\text{pres}$ is abstracted over, is well-formed and yields the exact same meaning as (36a).

\[
(40) \quad \exists [\lambda \alpha_2 [\text{John water-past}^{\text{ENG}}_0,2 \text{the plant}]][\text{before}^{\text{B&C}}_0 [\lambda \alpha_0 [\exists [\lambda \alpha_3 [\text{die-past}^{\text{ENG}}_0,3]]]]]
\]

\(^7\) There are cases of acceptable English present-under-past, such as $\text{I watered the plant before it dies on me}$, uttered in a situation where there is a good chance that the plant will die tomorrow. We briefly discuss these cases in 6.2.

\(^8\) It should be noted that Japanese does not clearly license NPIs in $\text{before}$-clauses. Kimiko Nakanishi (p.c.) points out (i) – an almost but far-from-perfectly acceptable example, found on the web, with a minimizing NPI.

\[
(i) \quad \text{Buturi-no siken-wa [yubi ip-pon ugokas-u maeni] owat-ta.}
\]

\[
\text{physics-GEN exam-TOP [finger one-CL move-PRES before] end-PAST}
\]

\`
\text{The physics exam ended before I moved/lifted a finger.}'
\`

This may be taken to mean that an Oghihara-style before (see (14); see also Kubota et al. 2011) might work for Japanese just as well as a B&C-$\text{before}$. However, the behavior, typology and distribution of Japanese NPIs are, as of yet, not so well understood, and in any event they differ quite a bit from the behavior, distribution and typology of English NPIs (see Nakanishi 2006 and Shimoyama 2011). Therefore, a downward entailing $\text{before}$ is not obviously inconsistent with the Japanese NPI facts. Since our goal is to offer a theory that does not posit different before’s for different languages, we assume that Japanese, like English, has the B&C-style downward entailing before.
This is obviously an unwelcome result, and we may avoid it by stipulating that the English present has the meaning in (41) – where ‘o’ stands for ‘overlap’ – rather than (39) (the Japanese present as defined in (38) does not change).

\[(41) \ [[\pres^\text{ENG}_{k,j}]^o] \text{ is defined only if } g(j) \circ g(k). \text{ Whenever defined, } [[\pres^\text{ENG}_{k,j}]^o] := g(j)\]

The ‘o’-semantics of \(\pres^\text{ENG}\) solves the problem: as shown in (42) (cf. (25)), we now get IPF – with present- as well as past-under-past – whenever the evaluation index is \(\lambda\)-bound by \(\text{before}\), whether its referential index is \(\exists\)-closed or left free.

\[(42) \text{ a. } \exists [[\lambda_2 [\text{John water-past}_{0,2} \text{ the plant}]][\text{before}_{B&C} [\lambda_0 [\exists [\lambda_3 [\text{it die-past}_{0,3}]]]]] \text{ EARLIEST}_{C}(\{t\ | \text{there is a } t' > t \text{ such that the plant dies at } t'\}) = \text{MIN}(C) \text{ or is undefined}\]

\text{ b. } \exists [[\lambda_2 [\text{John water-past}_{0,2} \text{ the plant}]][\text{before}_{B&C} [\lambda_0 [\exists [\lambda_3 [\text{it die-past}_{0,3}]]]]] \text{ EARLIEST}_{C}(\{t\ | \text{there is a } t' < t \text{ such that the plant dies at } t'\}) = \text{MIN}(C) \text{ or is undefined.}\]

\text{ c. } \exists [[\lambda_2 [\text{John water-past}_{0,2} \text{ the plant}]][\text{before}_{B&C} [\lambda_0 [\exists [\lambda_3 [\text{it die-past}_{0,3}, \text{past}_{0,3}]]]] \text{ EARLIEST}_{C}(\{t\ | \text{the plant dies at } g(3)\}) = \text{MIN}(C) \text{ or is undefined.}\]

Secondly, although we assume that \(\exists\)-closure is optional, we must assume that \(\exists\) cannot bind the evaluation index of a pronominal tense (regardless of \text{before}). If we allowed it, we would effectively predict there to be no past/present contrast all, as shown in (43).

\[(43) \text{ a. } \{t\ | \text{there is a } t' > t \text{ such that the plant dies at } t'\} = \{t\ | \text{the plant dies at } t\}\]

\text{ b. } \{t\ | \text{there is a } t' < t \text{ such that the plant dies at } t'\} = \{t\ | \text{the plant dies at } t\}\]

Hence: \( [[\lambda_3 [\exists [\lambda_0 [\text{the plant die-past}_{0,3}]]] = [[\lambda_3 [\exists [\lambda_0 [\text{the plant die-pres}_{0,3}]]]]\]

We assume, then, that the prohibition against binding of an evaluation index by \(\exists\) is universal. (37a) is well-formed as far as indexing is concerned (the referential index is \(\lambda\)-bound; the evaluation index is \(\lambda/\exists\)-free), but it still leads to the Stump effect (or something very close to it), even under (41): although (37a) is now interpreted as making a more informative assertion than (37b), it still asserts something that follows from the meaning of the tenses themselves, namely, that the past watering precedes the end of the SP-overlapping dying. Therefore, (37a) is still ruled out as a pragmatic oddity.

It seems that even with the ‘o’-semantics of \(\pres\) in (41), not all unwanted LFs are ruled out by IPF (or by the requirement that \(\exists\) bind only referential indices), and some additional indexing stipulations are needed. To see this, consider (44), where \(\pres\) carries two identical indices. It is a conceivable LF, but nothing so far rules it out, since the presupposition contributed by the present tense in (44) is trivially...
satisfied: \{t \circ t \mid \text{the plant dies at } t\} = \{t \mid \text{the plant dies at } t\}, and IPF is circumvented.

(44) \[ \exists [\lambda_2 [\text{John water-past}_0^1,2 \text{ the plant}]] \text{before}^{\text{wcf}} [\lambda_3 [\text{it die-pres}_3^0,3,3]]] \]

We rule out this LF by stipulating that the evaluation index of the English present has to be free.

Do we need to stipulate both that English pres has an ‘o’-semantics and that its evaluation index has to be free? It appears that we do, for the following reason. As shown in Section 1, and illustrated in (45), Japanese can express simultaneous readings of attitude reports with an embedded present tense. But this isn’t the case in English, as shown by (46), which does not have a “simultaneous” reading, only a “double access” reading (requiring that Mary’s sickness overlap the finding out time, which precedes SP, as well as SP itself).

(45) Taro-o-top Hanako-no-nom be-sick-pres that say-past
Taro said: “Hanako is sick”

(46) John found out that Mary is sick.

This means that (47) is a good LF for the Japanese sentence (because Japanese PRES has an ‘=’-semantics; see (38)), but we have to exclude (48a) as well as (48b) as LFs of the English sentence. They are indeed excluded by the stipulation that requires the evaluation index of the present tense to be free. In fact, the only well-formed LF of (46) is (49), where the embedded present is completely free and interpreted ‘de re’, as overlapping the finding out time as well as SP (see Heim 1994 and Abusch 1997 for details).

(47) \[ t_0 \text{PAST}_0^1 [\lambda_1 [\text{Taro say-t }] [\text{PRES}_0^1 [\lambda_2 [\text{Hanako be-t }_2^3,3 \text{ sick}]]]]] \]

(48) a. \[ \exists [\lambda_2 [\text{John find out-past}_0^0,2 [\lambda_0 [\text{Mary be-pres}_0^0,0,0 \text{ sick}]]]]] \]

b. \[ \exists [\lambda_2 [\text{John find out-past}_0^0,2 [\lambda_0 [\exists [\lambda_3 [\text{Mary be-pres}_0^0,3,3 \text{ sick}]]]]]] \]

(49) \[ \exists [\lambda_2 [\text{John find out-past}_0^0,2 [\lambda_3 [\text{Mary be-pres}_0^0,5,5 \text{ sick}]]]] \]

The special “double access” interpretation of (49) depends on the ‘o’-semantics in (41), which allows the sickness time to stretch from SP all the way back to the time of finding out.

\[ \text{Say takes an argument of type } <s, <i, t>> \text{, so it combines with its syntactic complement – which is of type } <i, t> – \text{ by Intensional Functional Application (see Heim & Kratzer 1998 and Fn. 15).} \]

\[ (i) \text{ If } [\lambda w’ \in W. [\alpha \beta]^{C,w’} \in \text{Dom}( [\alpha]^{C,w’})], \text{ then } [\alpha \beta]^{C,w’} = [\alpha]^{C,w’}([\lambda w’ \in W. [\beta]^{C,w’}).] \]
3.3 (No-)p-shiftability of before-clauses

As some of the readers probably noticed, the meaning in (35b) (of the well-formed English John watered the plant before it died) and the meaning in (36b) (of its Japanese well-formed present-under-past counterpart) are not quite the same. The former restricts the dying time to a time preceding SP; the latter is compatible with the dying being before SP, concurrent with SP, or after SP. The question is whether this predicted contrast is justified.

It turns out that it is. Japanese exhibits a property we call p-shiftability, whereas English exhibits a property we call No-p-shiftability (see also Kaufmann & Miyachi 2008; Kubota et al. 2011). That English exhibits No-p-shiftability is illustrated in (50): both answers in (50a) are felicitous, but the second answer in (50b) sounds contradictory (in particular, in contexts where John’s leaving and Sally’s arrival are completely independent of each other).

(50) Bill: When did John leave? Did Sally already arrive?
   a. Fred: John left before Sally’s arrival, which took place this morning / which is scheduled to take place tomorrow.
   b. Fred: John left before Sally arrived, which she did this morning / #which she will tomorrow.

The Japanese (51), on the other hand, is felicitous, and means “Mary is scheduled to arrive tomorrow; so it is true that John departed before Mary’s arrival”.

(51) Mary-wa asu toochakusu-ru koto-ni nat-te i-ru.
   Mary-TOP tomorrow arrive-PRES fact-DAT become-GER be-PRES
   dakara, John-ga Mary-ga toochakusu-ru maeni syuppatsusi-ta to
   So John-NOM Mary-NOM arrive-PRES before depart-PAST say fact-
   iu koto-wa hontoo-da.
   TOP true

It is important to distinguish No-p-factivity, which obtains in both English and Japanese, from No-p-shiftability. No-p-factivity was illustrated in Section 2 by the fact that (13a) – John watered the plant before it died – does not entail that the plant dies prior to SP (or at all; the same applies to the Japanese (28)). We attributed No-p-factivity to the intensionality of before, and this raises the question of whether p-shiftability in Japanese is simply a by-product of the intensionality of before. The answer seems to be that it is not, for the following reason. An “intensional” interpretation of ‘q before p’ is one where the actual world itself
need not be among the worlds accessible from the actual world at the q-time, and therefore there need not be a p-time in the actual world (though nothing prevents the existence of such a time; see (18) and Appendix). It turns out (and this is not reflected in (18)), that an “intensional” interpretation requires that there be some causal connection between q and p. A causal connection is established either when q may prevent p (as in Clint died before he met his grandchildren, which does not entail that Clint actually met his grandchildren prior to SP; see Heinämäki 1974; Ogihara 1995), or when q may prevent some (bad/unwanted) consequence of p (as in I am glad Mary came back before it got dark when uttered before sunset, which does not entail that it actually gets dark prior to SP; Roger Levy, p.c.). When there is no plausible causal connection between q and p (as in The price of tomatoes went up before Mary finished her dissertation), ‘q before p’ is biased towards an “extensional” interpretation, according to which the actual world itself IS among the worlds accessible from the actual world at the q-time, with the result that there is a p-time in the actual world (in our case, Mary finishes her dissertation prior to SP).

Whatever the explanation of the effect of the causal connection on p-factivity might be (see Ogihara 1995 for some suggestions), it is quite clear that No-p-shiftability is a different effect. Indeed, the which-she-will-variant of Fred’s response in (50b) may be felicitous in contexts where there is a causal connection between John’s leaving and Sally’s arrival. But in contexts where John’s leaving has no effect whatsoever on Sally’s arrival (suppose they don’t know each other, have never heard of each other, their jobs and travel plans have nothing to do with each other, and so on), that variant of Fred’s response is infelicitous. On the other hand, the Japanese (51) may be felicitous in such contexts. This means that No-p-shiftability cannot be reduced to No-p-factivity. For if it could be reduced to that, we would not expect there to be a contrast between the English (50b) and the Japanese (51). We conclude that since a p-factive reading of a past-under-past before-clause entails No-p-shiftability in English, but a p-factive reading of a present-under-past before-clause does not entail No-p-shiftability in Japanese, (No-) p-shiftability is independent of the intensionality of beforeB&C. In sum, the LFs in (35a) and (36a) are compatible with the observation that Japanese exhibits p-shiftability whereas English exhibits No-p-shiftability.

We now turn to the Polish facts. As we saw in Section 1, Polish follows the Japanese pattern in attitude reports and the English pattern in before-clauses. So first we have to examine the compatibility of our proposal with the behavior of English and Japanese tenses in attitude
4 POLISH AS A HYBRID LANGUAGE

4.1 SOT and non-SOT languages

English is considered to be an SOT language, because sometimes – for example, in attitude reports – a past tense embedded under another, c-commanding, past tense can be “invisible” to semantic interpretation, presumably, as a result of the application of the SOT-rule, which can be informally stated as follows: “Optionally delete a tense that is c-commanded by a morphologically agreeing tense”. This suggests the following alternative to the theory of before outlined in Section 3, namely, that the past tense in English before-clauses may be “invisible” to semantic interpretation as well. Let us explore this theoretical option.

The motivation for assuming an SOT rule in English is illustrated in (52), which has a “simultaneous” reading, according to which Mary’s sickness occurs during John’s “now”, in addition to a “back-shifted” reading, according to which Mary’s sickness precedes John’s “now”.

(52) John said that Mary was sick.
   a. John: “Mary is sick” (Simultaneous reading)
   b. John: “Mary was sick” (Back-shifted reading)

As we saw in Section 1, the Japanese past-under-past counterpart of (52) – namely, (8) – has only the back-shifted reading.

To account for this, we do not have to assume (as we did in 3.2) that English and Japanese tenses differ in semantic type. In fact, we may assume that: (i) the past tense in both English and Japanese is quantificational (see (53)); (ii) English has the SOT rule whereas Japanese does not; and (iii) a “deleted” English past tense, $\text{PAST}$, has no meaning. This would allow us to posit the LFs in (54) and (55) (see von Stechow 1995, 2002; Ogihara 1996 among others). Recall that on our assumptions, a quantificational past tense must QR, but the same predictions would be made if the past were base-generated in a pre-sentential position.

(53) $[[\text{PAST}^\text{ENG}]^\text{C,B}] = [[\text{PAST}^\text{JAP}]^\text{C,B}}$

(54) $[t_0 \ [\text{PAST}^\text{ENG} \ [\lambda_1 \ [\text{John say-t}_1 \ [\\text{PAST} \ [\lambda_2 \ [\text{Mary be-t}_2 \ \text{sick}]]]]]]] \quad \text{“simultaneous”}$

(55) $[t_0 \ [\text{PAST}^\text{ENG/JAP} \ [\lambda_1 \ [\text{John/Taro say-t}_1 \ [\\text{PAST}^\text{ENG/JAP} \ [\lambda_2 \ [\text{Mary/Hanako be-t}_2 \ \text{sick}]]]]]]] \quad \text{“back-shifted”}$
Can we make similar assumptions regarding before-clauses, as suggested by Ogihara (1996)? The Japanese past-under-past LF would still be (34a) (and ruled out by IPF); the English past-under-past LF would be (56).

\[(56) \quad [t_0 \left[PAST^{ENG} [\lambda_1 [John \text{ water-t}_1 \text{ the plant}]] \left[before^{B&C} [PAST [\lambda_2 [it \text{ die-t}_2]]] \right]]\right] \]

(56) avoids IPF because \[[\left[before^{B&C} [PAST [\lambda_2 [it \text{ die-t}_2]]] \right]]^{C-g} = \left[[before^{B&C} [\lambda_2 [it \text{ die-t}_2]] \right]]^{C-g}, \] and earliest picks out the leftmost dying moment. But clearly the SOT analysis makes wrong predictions: (56) implies that the dying can occur during or after SP, but as we saw in 3.3, English before-clauses exhibit No-p-shiftability. Therefore, we maintain the pronominal analysis outlined in 3.2 for English, repeated below as (57), which predicts No-p-shiftability.

\[(57) \quad \exists \left[[\lambda_3 [John \text{ water-past}^{ENG}_{0,3} \text{ the plant}]] \left[before^{B&C} [\lambda_2 [it \text{ die-past}^{ENG}_{0,2}]] \right]\right] \]

This pronominal analysis is compatible with the SOT analysis of English attitude reports, as long as we assume that pronominal tenses can undergo “deletion” by the SOT-rule, resulting in loss of one of the two indices. A “deleted” past is interpreted as in (58), and we preserve the predictions about past-under-past in English attitude reports.

\[(58) \quad \left[\text{past}^{k}\right]^{g} := g(k). \]

\[(59) \quad \begin{align*}
\text{a.} & \quad \exists \left[\lambda_2 [John \text{ say-past}^{ENG}_{0,2} [\lambda_0 \left[\exists \lambda_3 [Mary \text{ be-past}^{ENG}_{0,3} \text{ sick}]]\right]]\right] \quad \text{“back-shifted”} \\
\text{b.} & \quad \exists \left[\lambda_2 [John \text{ say-past}^{ENG}_{0,2} [\lambda_3 [Mary \text{ be-past}^{ENG}_{0,3} \text{ sick}]]\right] \quad \text{“simultaneous”} \\
\end{align*} \]

Crucially, we assume that the fact that \[\text{say-past}^{ENG}_{0,2}\] c-commands the embedded past in (59b) suffices to trigger the SOT-rule. By contrast, the SOT-rule is inapplicable in (60), unlike (56), because \[\text{water-past}^{ENG}_{0,3}\] does not c-command the embedded past.

\[(60) \quad \exists \left[\lambda_3 [John \text{ water-past}^{ENG}_{0,3} \text{ the plant}]] \left[before^{B&C} [\lambda_2 [it \text{ died-past}^{ ENG}_{2}]] \right]\right] \]

So far we have assumed that while quantificational tenses must QR (to avoid type-mismatch), pronominal tenses need not. Now we have to assume that pronominal tenses cannot QR: if they could, we would wrongly predict there to be another LF in addition to (60), where the matrix past QRs to a position c-commanding the embedded past, triggering its “deletion” (and leading to p-shiftability). Since we assume that QR of pronominal tenses is not possible, English is correctly predicted to exhibit No-p-shiftability, despite the fact that it is an SOT language. The only well-formed past-under-past before-clause is one that looks like (35), which implies No-p-shiftability. However, as we will see in Section 6, in SOT languages with quantificational tenses, the matrix past
can trigger deletion of the past in the before-clause (giving rise to p-shiftability).

The general picture that emerges regarding tense contrasts between English and Japanese is summarized in (61).

If we assumed that Japanese had pronominal tenses, in particular, a pronominal past (instead of, or in addition to, a quantificational past), we would wrongly predict Japanese to have past-under-past before-clauses (like English). On the other hand, if we assumed that English had quantificational tenses, we would predict it to exhibit p-shiftability in before-clauses (as a result of the matrix past QR-ing and “deleting” the embedded past).

The SOT rule applies under agreement, which means that it applies to present-under-present, just like it applies to past-under-past. We claimed in 3.2 that the evaluation index of the English present has to be free. This constraint does not apply when the present is “deleted” by the SOT rule (in which case the pronoun has just one index) as suggested by (62), which need not have a “double access” reading: Mary’s sickness can be in the future.

(62) John will find out that Mary is sick.

This supports the view (e.g., Abusch 1997) that will is composed of the modal woll and the present tense; the latter triggers “deletion” of the embedded present.

The assumption that the SOT rule applies to preset-under-present/future also explains Stump’s (1985) observation that a present-under-future is well-formed in English and crucially refers to a time following SP rather than SP itself.

(63) John will water the plant before it dies.

As pointed out by Oghihara (1996) and von Stechow (2009), (63) looks like a case of tense “deletion”, just like (62), and just like John will catch a fish that is alive (which has a reading according to which the fish is alive during the catching time). To bring out this similarity in our system, we
assume that *woll*, which is an existential quantifier over future times, QR’s together with its argument – the present tense (see Ogihara & Sharvit 2012). This movement creates the necessary configuration for the SOT-rule to apply (cf. (59b), where the past tense that is “glued” to *say* triggers “deletion” of the embedded past). 10

(64) a. *woll-pres$_{0,3}$* [[λ$_1$ [John water-t$_1$ the plant]] [before$^{B&K}$ [λ$_5$ [it die-pres$_3$]]]]

b. *woll-pres$_{0,3}$* [λ$_1$ [John catch-t$_1$ a fish [wh$_1$ e$_5$ be-pres$_1$ alive]]]

c. *woll-pres$_{0,3}$* [λ$_2$ [John find-out-t$_2$ [λ$_3$ [Mary be-pres$_3$ sick]]]]

The SOT rule is not obligatory, so for *John will catch a fish that is alive* and *John will find out that Mary is sick* we get another reading (derived from an LF where the embedded present is not “deleted” and overlaps SP), but for (63), (64a) is the only semantically coherent LF, as failure to “delete” the embedded present results in a semantic anomaly (according to which a future time precedes SP). 11 The SOT theory also explains the “simultaneous” readings of *Clean your room before your guests arrive* and *Mary always cleans her room before her guests arrive*, by positing a covert present tense with widest scope, attached to some (overt or covert) quantifier, triggering “deletion” of embedded present tenses.

(65) **IMPERATIVE-pres$_{0,1}$** [[λ$_1$ [clean-pres$_1$ your room]] [before [λ$_3$ [your guests arrive-pres$_3$]]]]

**ALWAYS-pres$_{0,1}$** [[λ$_5$ [Mary clean-pres$_5$ her room]] [before [λ$_3$ [her guests arrive-pres$_3$]]]]

Some clarification is in order, regarding the application of the SOT rule in past-under-past configurations. It is sometimes claimed in the literature that certain languages that are traditionally classified as non-SOT languages have “simultaneous” readings of past-under-past in attitude reports (see, for example, Althshuler 2008). This is not predicted by the theory advocated here, as “simultaneous” readings of past-under-past require the SOT rule. Indeed, as it turns out, languages that are classified as non-SOT have, what look like, “simultaneous” readings of past-under-past only in a proper subset of the attitude environments where “pure” SOT languages allow “simultaneous”

---

10 Stump’s other observation, that *will-under-will before-clauses are ungrammatical (*John will leave before Mary will arrive)*, is accounted for by IPF, just like past-under-past, because the EARLIEST-expression fails to refer (see Sharvit & Tieu 2011). However, this analysis (as other analyses we are aware of) faces problems in view of counter-examples to Stump’s observation, such as the well-formed (i) (Ogihara 1996) and (ii) (Carson Schütze, p.c.).

(ii) Mubarak must resign before the crowd will lose control/loses control.

11 One might wonder, however, why tense “deletion” is impossible in *(John was still happy then, but) Mary would leave him before he realized it* (i.e., why it cannot mean that John’s realization is after SP). The reason probably has to do with the special pragmatics of “non-deleted” *would*, which is felicitous only if by replacing it with the past tense (in our case, *Mary later left him before he realized it*) preserves the intended meaning (see Fn. 14 for a related problem).

(i) John will leave before Mary will/does.

(ii) Mubarak must resign before the crowd will lose control/loses control.
readings. For example, while it is true that some languages classified as non-SOT indeed have (what appears to be) a “simultaneous” reading of the counterpart of (66a), according to which John’s utterance is: “Mary is sick”, none of these languages has a “simultaneous” reading of the counterpart of (66b) which, according to the theory advocated here, has the LF in (67), where the matrix past tense triggers “deletion” of the embedded past tenses, including the past-argument of woll (implying that John’s utterance was: “Mary will arrive before I leave”, and neither arrival nor leaving are confined to times preceding SP).

(66) a. John said that Mary was sick.
   b. John said that [Mary would arrive before he left]

(67) \[ \text{John say-past}_\text{ENG}^0,6 \lambda_3 [\text{woll-past}_\text{ENG}^0,6 \{ [\lambda_1 \text{[Mary arrive-t]}_\text{ENG}^0,2] \text{[before B&C]} [\lambda_5 \text{[he leave-past]}_\text{ENG}^0,5] \}] \]

Therefore, it makes sense (as Grønn & von Stechow 2010 and Ogihara & Sharvit 2012 do) to say that there are “pure” SOT languages where “simultaneous” readings are the result of the application of the SOT-rule. “Pseudo-simultaneous” readings (say, of the counterpart of (66a)) in non-SOT languages result from a ‘de re’ interpretation of the embedded past. Given this, it comes as a surprise that for many speakers, (68a) implies that the expected arrival precedes SP. Perhaps for those speakers the infinitival clause blocks “deletion” (on the locality of “deletion”, see Schlenker 1999 a.o.). Borrowing Grønn & von Stechow’s (2011) idea that dies in (64a) contains a hidden future (rather than a “deleted” present) would also explain why for many speakers (68b) can imply that the expected arrival follows SP (this pronominal future is, presumably, subject to the same indexing constraint as pres\text{\textit{ENG}}, see 3.2).

(68) a. My parents told me to clean my room before my guests arrived.
   \[ \text{... tell-past}_\text{ENG}^0,6 \lambda_3 \{ [\text{before B&C} \{ [\lambda_2 \{ \text{arrive-past}_\text{ENG}^0,6 \} \}] \} \]
   b. My parents told me to clean my room before my guests arrive.
   \[ \text{... tell-past}_\text{ENG}^0,6 \lambda_3 \{ [\text{before B&C} \{ [\lambda_2 \{ \text{arrive-fut}_\text{ENG}^0,6 \} \}] \} \]

Having expanded our theory to cover attitude reports in addition to before-clauses, we are now ready to discuss the Polish pattern.

---

12 A ‘de re’ interpretation of the past on left in (66b) is excluded by the Upper Limit Constraint (Abusch 1997), which does not allow John to use a description such as “tomorrow” to describe to himself a time that precedes SP but follows his saying time (and in any event would incorrectly restrict the leaving to times preceding SP). The LFs in (68) are consistent with this because, presumably, the Upper Limit Constraint applies only to referential indices of tenses (and only when they are free). In addition, since we assume that Japanese does not have pronominal tenses, we predict the Japanese counterpart of (66a) not to have a "(pseudo-)simultaneous" reading (as observed in Ogihara & Sharvit 2012).
4.2 The Polish pattern and the present tense cross-linguistically

As shown in Section 1, Polish exhibits the English pattern in before-clauses. Furthermore, Polish past-under-past before-clauses exhibit No-p-shiftability (Marcin’s coming in (5) must precede SP). So it may seem that all we need to say is that Polish has pronominal tenses, thus predicting the well-formedness of past-under-past before-clauses and the ill-formedness of present-under-past before-clauses. Since Polish also, presumably, lacks the SOT rule, we also correctly predict that in attitude reports past-under-past does not have a simultaneous reading. Such a theory, however, will yield correct results only for past-under-past: we should also predict that the Polish present exhibits an English-like behavior in before-clauses, and a Japanese-like behavior in attitude reports (as evidenced by (9)). So if Polish indeed has pronominal tenses, the theory somehow has to allow its pronominal present tense to differ from the English pronominal present tense.

We suggest, then, that the semantics of Polish pres is the same as that of English (see (41)). But in addition to differing with respect to the SOT rule, the two languages also differ with respect to the syntax of pres: English has a rule that requires its evaluation index to be free;\(^\text{13}\) Polish obeys a different indexing constraint: both its indices may be bound, but not by the same binder. As a result, Polish – as opposed to English – can have (69a) (where the embedded pres is locally existentially closed) as a well-formed attitude LF, but the before-LF in (69b) (where the embedded present is locally existentially closed) is still ruled out as a presupposition failure.

\(^{13}\) Daniel Alshuler (p.c.) and Seth Cable (p.c.) point out that (i) has, for many speakers, a reading that implies that John’s future utterance is “We are having our last meal together now”, much like Abusch’s (1997) (ii) (where both embedded past tenses may be “deleted”), in addition to a reading that implies that John’s future utterance is “We are having our last meal together in a little while”.

(i) A week ago, John decided that in ten days he would tell his mother that they are having their last meal together.
(ii) A week ago, John decided that in ten days he would tell his mother that they were having their last meal together.

We suggest that for these speakers, the restriction on the English present is that it has to be free in the scope of the closest c-commanding tense-binding operator, making [John decide-past,3 \[\lambda_5 \mbox{well-}pres_5 [\lambda_0 \exists [\lambda_4 \mbox{tell-60 his mother } \lambda_8 \mbox{they be-pres_0,4 have...}]...]]] a possible LF, where the embedded present is interpreted ‘de re’ in relation to tell (making it possible for the meal-having to be concurrent with the telling) but it is a bound variable in relation to decide and will. This implies that it should be possible to interpret the most deeply embedded past in (ii) ‘de re’ relative to tell. But there is no pragmatically plausible acquaintance-based description applied to the past tense that would yield John’s “now” in the worlds compatible with what he tells his mother and the time of telling in the future worlds compatible with his decision (unless John decided, a week ago, to attribute to himself a future misidentification of the time in which he will be located).
(69)  
\[
\begin{align*}
\text{a. } & \exists [\lambda_3 \text{[Ania say-past}_{0.3}^\text{POL} \ [\lambda_2 \exists [\lambda_5 \text{[Marcin be-pres}_{2.5}^\text{POL} \text{ sick}]]]]] \\
\text{Semantically and syntactically well-formed}
\end{align*}
\]

\[
\begin{align*}
\text{b. } & \exists [[\lambda_2 \text{[John water-past}_{0.2}^\text{POL} \text{ the plant}]] \text{[before}_{B&C}^\text{POL} \ [\lambda_0 \exists [\lambda_3 \text{[it die-pres}_{0.3}^\text{POL}]]]]] \\
\text{Semantically ill-formed (cf. (42)).}
\end{align*}
\]

On the other hand, like English, Polish cannot have (70a,b) as possible LFs, because the two indices of pres are bound by the same binder.

(70)  
\[
\begin{align*}
\text{a. } & \exists [\lambda_3 \text{[Ania say-past}_{0.3}^\text{POL} \ [\lambda_2 \exists [\lambda_5 \text{[Marcin be-pres}_{2.2}^\text{POL} \text{ sick}]]]]] \\
\text{b. } & \exists [[\lambda_2 \text{[John water-past}_{0.2}^\text{POL} \text{ the plant}]] \text{[before}_{B&C}^\text{POL} \ [\lambda_0 \exists [\lambda_3 \text{[it die-pres}_{0.3}^\text{POL}]]]]]
\end{align*}
\]

This correctly predicts a contrast between Polish attitude reports and Polish before-clauses, despite the fact that Polish has a “shiftable” present tense in attitude reports. (71) summarizes the cross-linguistic assumptions (recall that when free, the evaluation index of a pronominal tense denotes SP).

(71) Japanese Present: Quantificational; fully shiftable.
Polish Present: Pronominal; semi-shiftable (i.e., can be bound, though not by the same binder that binds the referential index).
English Present: Pronominal; non-shiftable (i.e., must be free).

Independent motivation for the pronominal/quantificational tense distinction comes from the contrast between Polish/English and Japanese relative clauses (see Ogihara & Sharvit 2012 for related observations): in the Polish (72), with present-under-past, the standing time is forced to overlap SP (as in English: John met a woman who is standing by the tree); in the Japanese (73), the travel-loving time need not overlap SP (as opposed to the English Joseph met a woman who loves traveling where the travel-loving time must overlap SP).

(72) Jan spotka kobietę która stoi przy drzewie.
Jan meet-PAST woman-ACC which.NOM stand-PRES by tree

(73) kodomo-no koro, Joseph-wa ryokoo-o aisuru zyosei-ni at-ta
child-GEN time, Joseph-TOP[travelling-ACC love-PRES woman]-DAT meet-PAST

Why shouldn’t the Polish present be “null” in relative clauses, given that it is “null” in attitude reports? The assumption that Polish is indeed non-SOT, but unlike Japanese it has pronominal tenses, as defined in (39), provides an explanation: the Japanese past tense is quantificational, undergoes QR, as in (74), and binds the evaluation time of the embedded present. This option is not available in Polish or English: the evaluation index of the embedded present in (75) remains free and refers to (a time that overlaps) SP.
On the other hand, past-under-past relative clauses (e.g., John met a woman who loved travelling) are possible in all three languages (and the two past tenses can accidentally “corefer”; see von Stechow 1995 for an argument against tense “deletion” in relative clauses). 14

It follows from these assumptions that a language can have a “null” present in both attitude reports and relative clauses only if its present is quantificational. If its present is pronominal, it can be “null” in attitude reports, but not in relative clauses (as implied by (74)–(75)), because to satisfy the c-command requirement of the SOT-rule, QR must take place, and pronouns (by assumption) do not QR.

We have to be careful, from now on, about the use of the term “null” (because Polish has a “null” present in attitude reports but not in before-clauses). Let us, then, agree on the following terminology: the term “null present” refers to how the present is understood, and is therefore construction-specific – in a given language, the present can be “null” in one construction but not in another (Polish is such an example). The term “shiftable present” on the other hand, is not construction-specific, and refers to the lexical properties of the item in question (see (71)); in that sense, both Japanese and Polish have a “shiftable present”, but in Polish the present is only “semi-shiftable”, because it obeys an indexing constraint which Japanese does not. English does not have a shiftable present tense at all because English obeys a stricter indexing constraint.

Interestingly, it follows from (71) that any language with a semi-shiftable quantificational present tense (call it PRESXXX) gives rise to an LF that has the Stump effect and/or IPF; see (77).

14 We assume that pronominal tenses never QR, not even in John met the woman who would hire him (where would is composed of a past and woll and) where the hiring cannot precede the meeting. Indeed, if we assumed that the embedded past is “deleted”, we would account for this, but notice that we would still need a pragmatic explanation for why the hiring cannot occur after SP (cf. Heim 1994, who attributes this observation to J. Higginbothm; see Fn. 11). Also, see von Stechow (1995) for arguments against tense “deletion” in English relative clauses.
That is to say, we do not expect to find a language whose present tense exhibits a Japanese-like behavior in both relative clauses and attitude reports and a Polish/English-like behavior in before-clauses. Furthermore, we predict that there cannot be a language with a well-formed present-under-past in before-clauses and a “null” present-under-past in relative clauses (i.e., a Japanese-like present), but no “null” present-under-past in attitude reports. The reason is that a Japanese-like present is fully-shiftable; as such it must support “simultaneous” readings in attitude reports. The existence of a language with a Japanese-like present in before-clauses and relative clauses but an English-like present in attitude reports would falsify the theory of indexing constraints (i.e., the distinction between full- and semi-shiftability), because it would show that there is no distinction between “null” and “shiftable” – both are construction-specific. (78) summarizes the predictions (P = pronominal; Q = quantificational).

(78) | Present | P/Q | before-clauses | attitude reports | rel. clauses |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Non-shiftable</td>
<td>P</td>
<td>* (Stump effect (37))</td>
<td>√, only “double access” ((49))</td>
</tr>
<tr>
<td>Polish</td>
<td>Semi-shiftable</td>
<td>P</td>
<td>* ((70b)/IPF (69b))</td>
<td>√, “simultaneous” ((69a))</td>
</tr>
<tr>
<td>Japanese</td>
<td>Fully shiftable</td>
<td>Q</td>
<td>√ ((36))</td>
<td>√, “simultaneous” ((47))</td>
</tr>
</tbody>
</table>

As (78) implies, the only reason to say that Japanese has a quantificational present (rather than a fully shiftable pronominal present) is its behavior in relative clauses: if the Japanese present were a pronoun, it wouldn’t QR, leading to an English/Polish-like behavior in relative clauses.

In Section 5 we consider and reject an alternative analysis of before, the relative clause analysis, which avoids IPF in past-under-past before-clauses without appealing to the pronominal/quantificational tense distinction. This alternative analysis does not depend on the assumption that the time axis is dense and, as such, it poses a potential challenge for the main claim of the current proposal (namely, only pronominal tenses or the application of the SOT rule can avoid IPF). Readers who wish to skip Section 5 may go directly to Section 6, where the broader cross-linguistic implications of the pronominal/quantificational tense distinction are discussed.
5 GEIS-AMBIGUITIES AND THE RELATIVE CLAUSE
ANALYSIS OF BEFORE-CLAUSES

Recall our main claim that no language can have a quantificational past tense in a before-clause, due to IPF, as shown again in (79).

(79) a. \( \text{before}^{B&C} [\text{PAST} [\lambda_2 [\text{it die-t}_2]]] \)

b. \( \text{EARLIEST}_C(\{t \mid \text{there is a } t' < t \text{ such that the plant dies at } t'\}) \) is undefined.

There is an alternative analysis (e.g., von Stechow 2009, based on Heim 1997) where IPF is avoided even if the past tense is taken to be quantificational (and without assuming that the SOT-rule applies in before-clauses). Adapting that analysis to our syntactic framework, the LF of the English \( \text{John watered the plant before it died} \) may be the relative-clause-like (80a) where the semantically vacuous \( \text{wh}_3 \) is a silent relative pronoun, whose trace – \( e_3 \) – is a variable over \( D_i \); \( \text{at} \) introduces the identity relation (as in (80b)). \( \lambda_2 [\text{it die-t}_2] \) and \( [\text{at } e_3] \) – both of type \(<i, t>\) – combine by Predicate Modification.

(80) a. \( \text{before}^{B&C} \text{wh}_3 [\lambda_3 [t_0 [\text{PAST} [\lambda_2 [\text{it die-t}_2]] [\text{at } e_3]]]] \)

b. \( \lbrack \text{at} \rbrack^{C,g}(t')(t') = \text{True iff } t' = t. \)

c. \( \text{EARLIEST}_C(\{t \mid \text{there is a } t' < SP \text{ such that: the plant dies at } t' \text{ and } t' = t\}) = \text{EARLIEST}_C(\{t \mid t < SP \text{ and the plant dies at } t\}) \)

Movement of the relative pronoun \( \text{wh}_3 \) results in insertion of a \( \lambda \)-operator, and turns \( [t_0 [\text{PAST} [[\lambda_2 [\text{it die-t}_2]] [\text{at } e_3]]]] \) – which denotes an element of \( D_i \) – into \( [\lambda_3 [t_0 [\text{PAST} [[\lambda_2 [\text{it die-t}_2]] [\text{at } e_3]]]]] \), of type \(<i, t>\) (and an appropriate argument of \( \text{before}^{B&C} \)). The crucial difference between (79a) and (80a) is this. In (79a), \( \text{before}^{B&C} \) shifts the evaluation time of \( \text{PAST} \). Since \( \text{PAST} \) contributes an existential quantifier over times preceding the evaluation time, this results in IPF. In (80a), on the other hand, \( \text{before}^{B&C} \) does not shift the evaluation time of \( \text{PAST} \); instead, it binds the argument of \( \text{at} \) (which, by the semantics of \( \text{at} \), ends up being equal to the dying time). This results in neutralizing the existential, as shown in (80c). Accordingly, \( \text{EARLIEST} \) picks out the leftmost past dying time in \( C \), predicting No-p-shiftability.

This relative clause analysis also seems to come with the added bonus of predicting Geis-ambiguities (Geis 1970), as illustrated in (81): when the before-clause contains an intensional verb such as \( \text{say} \), \( [\text{at } t_3] \) can modify either an \(<i, t>\)-node “above” \( \text{say} \) (in this case, \( [\lambda_2 [\text{Mary say-t}_2 \text{ PAST} [\lambda_4 [\text{she arrive-t}_4]]]] \)), yielding a “high” reading, or an \(<i, t>\)-node “below” \( \text{say} \) (in this case, \( [\lambda_4 [\text{she arrive-t}_4]] \)), yielding a “low” reading.
(81) a. John watered the plant before Mary said she arrived.
    b. High reading: At 3pm, Mary said: “I arrived”. John watered the plant prior to Mary’s speaking time (3pm); i.e., prior to $\text{earliest}_C\{t \mid \exists t' < \text{SP} \text{ such that } \text{Mary says at } t' \text{ “I arrived” and } t' = t\}$. 
    \[
    \text{before}^{\text{B&C}}_{\text{wh}} [\lambda_3 [t_0 \text{PAST } [\lambda_2 [\text{Mary say}-t_2 \text{PAST } [\lambda_4 [\text{she arrive}-t_4]])]] [\text{at } e_3]]
    \]
    c. Low reading: Mary said: “I arrived at 1pm”. John watered the plant prior to Mary’s declared arrival time (1pm); i.e., prior to $\text{earliest}_C\{t \mid \text{Mary said “there is a past time } t' \text{ such that I arrived at } t' \text{ and } t = t'”\}$. 
    \[
    \text{before}^{\text{B&C}}_{\text{wh}} [\lambda_3 [t_0 \text{PAST } [\lambda_2 [\text{Mary say}-t_2 \text{PAST } [\lambda_4 [\text{she arrive}-t_4]]]] [\text{at } e_3]]]
    \]

The analysis is further motivated by the fact that in some languages (e.g., Russian; see Grønn & von Stechow 2011), the surface sister of before appears with an overt ‘at which’ relativizer. And even in English, John watered the plant before it bloomed may sometimes be paraphrased as John watered the plant before the time at which it bloomed, where at which it bloomed is clearly a relative clause. However, there are some facts that the analysis fails to account for.

First, (81c) implies, contrary to fact, that the low reading of (81a) entails that Mary is making a commitment about a specific time (or specific times) at which she arrived. While (81a) can certainly have such a “low” reading, it can have another “low” reading, according to which Mary might have said “I arrived whenever Grandma died”, while being non-committal about when exactly she herself arrived or Grandma died. On that construal, the inference is that John waters the plant prior to Grandma’s actual dying time – so if Grandma actually died at 3pm, we infer that John watered the plant prior to 3pm. But importantly, in contexts where Mary is mistaken about when Grandma died (for example, Mary is convinced Grandma’s death time is prior to 3pm), though she is well-informed about what time it is (crucially, Mary does not misidentify 3pm as some other time), she would not accept 3pm as Grandma’s dying time, or as her own arrival time, but (81c) does not reflect this. Furthermore, even if there was a sensible way to apply earliest to the relative clause (which obviously does not simply denote a set of times in this case, but rather a set of time-concepts), it wouldn’t necessarily yield the leftmost moment of Grandma’s actual dying time (again, because Mary does not accept 3pm as Grandma’s dying time, despite the fact that she correctly identifies 3pm).
Secondly, not all languages allow Geis-ambiguities in *before*-clauses lacking an overt *the time at which*, as shown by the Spanish (82)–(83).

(82) Juan llegó antes de que María pensara que Pedro se habia ido
Juan arrived before of that María would-think(subj) that Pedro cl. had left
‘Juan arrived before María thought Pedro left’ HIGH/*LOW

(83) Juan llegó antes del momento en el que María pensó
Juan arrived before of the moment in the which María thought(indc)
que Pedro se había ido
that Pedro cl. had left
‘Juan arrived before the moment in which María thought Pedro left’ HIGH/LOW

This seems to cut across the SOT/non-SOT distinction, as well as the distinction between languages that have past tense in *before*-clauses and languages that do not, because Spanish is SOT (see Section 6) and obviously allows past-under-past *before*-clauses; while Japanese, which is non-SOT and doesn’t allow past-under-past *before*-clauses, shows similar behavior regarding Geis-ambiguities (Arregui & Kusumoto 1998; Otaki 2010). This means that simply stipulating that Japanese *the-time-at-which*-less *before*-clauses do not have a relative-clause structure (Kusumoto 1999) may explain why Japanese *the-time-at-which*-less *before*-clauses do not have Geis-ambiguities, but cannot automatically explain why the past tense is not allowed there.

There are also differences between overt and covert *the time at which* even in English. As (84)–(88) show, an overt *the time at which* does not license NPIs; likewise, an NPI blocks any low reading of a sentence with a potential Geis ambiguity (regardless of whether the sentence contains an overt *the time at which*).

(84) John watered the plant before it (ever) bloomed.
(85) John watered the plant before the time at which it (*ever) bloomed.
(86) John watered the plant before Sally ever said it bloomed. HIGH/*LOW
(87) John watered the plant before Sally said it ever bloomed. HIGH/LOW
(88) John watered the plant before the time at which Sally (*ever) said
 it (*ever) bloomed. HIGH/LOW

These facts, taken together, suggest that the relative clause analysis cannot, on its own, tell the whole story of *before*.

As an alternative, we propose that *before*-clauses come in two varieties: there are “bare” *before*-clauses, and there are *before*-clauses preceded by *the time at which* (overt or covert). English allows a covert *the time at which*, but not all languages do (Spanish, for example, does not). Notice
that NPI-blocking in (85) and (88) is expected, given that *the* usually blocks NPIs by a “higher” licensor, as shown by the contrast in (89), where the licensor is negation (see Guerzoni & Sharvit 2007 for discussion of why *the* doesn’t license NPIs on its own).

(89) a. John likes no woman he ever dated.
   b. *John doesn’t like the woman he ever dated.

Given (89), we expect NPIs to be blocked when before is followed by a covert the time at which as well. We propose, then, that only before-clauses that are preceded by the time at which, overtly or covertly, produce low readings. This is why (86) and (87) lack a low reading: only the second before—the one that is accompanied by the time at which—produces low readings, in which case NPIs are blocked.

Inspired by Otaki (2010), we flesh out this idea as follows. We assume that there is a DP-taking before and a clause-taking before. The DP-taking before is before\textsuperscript{F} in (90), whose syntactic argument is a DP that may contain a relative clause with time as head and [at [...] ] as the relativization site.

(90) For any t,t′ ∈ D, \[before\textsuperscript{F}]^C\textsubscript{B&C}(t)(t′) := True iff t′ < t. \textit{ (cf. Heim 1997)}

Even if we extend the notion of entailment to cover intervals, so that an interval entails its super-intervals, rendering before\textsuperscript{F} downward entailing on its first argument, *the* blocks NPI-licensing by before\textsuperscript{F} just like it blocks NPI-licensing by not in (89b). The clause-taking before is before\textsuperscript{B&C} in (24) (or, more accurately, its intensional variant; see (18) and Appendix), which is downward entailing. Neither one of these before’s is language-specific, but languages may differ regarding the licensing of a covert the time at which.

We maintain our claim that English has pronominal tenses. In addition, a high reading may be derived with either before\textsuperscript{B&C} (see (91)) or before\textsuperscript{F} (see (92)); a low reading requires before\textsuperscript{F} (see (93)), which yields both ‘de re’ (specific) and ‘de dicto’ (non-specific) construals. The trace of the relative pronoun in (92) – e\textsubscript{5} – is of type i; and the trace of the relative pronoun in (93) – E\textsubscript{5} – is of type <<i, <s, i>> (denoting a function from time-world pairs to times). t\textsubscript{0} in (93), like t\textsubscript{0}, is a silent time-denoting pronoun. Time is semantically vacuous; its sister is a relative clause: of type <i, t> in (92); of type <<<i, <s, i>>, t> in (93). The is type-blind, and picks out the unique element (if there is one) of the set denoted by the relative clause. Accordingly, DP1 in (92) denotes an element of D\textsubscript{i}, and DP2 in (93) denotes an element...
of \( D_{i+s,i} \); in both (92) and (93), \([\text{before}_F]^{C,g}\) applies to an element of \( D_i \).

(91) High reading with \([\text{before}_B^F]^{C,g}\) – NPI OK (see (86)/(87))

a. \( \exists \left[ \lambda_3 \left[ \text{John water-past}^{\text{ENG},0,3 \text{ the plant}} \right] \left[ \text{before}_B^{C,g} \lambda_2 \left[ \text{Sally say-past}^{\text{ENG},0,2 \text{ it bloomed}} \right] \right] \right] \)

b. For all worlds \( w \) accessible from \( @ \) (= the actual world), the past watering time in \( @ \) precedes the leftmost past moment in \( C \) where Sally says in \( w \) “the plant bloomed”.

(92) High reading with \([\text{before}_F]^{C,g}\) blocked by the time (see (88))

a. \( \exists \left[ \lambda_3 \left[ \text{John water-past}^{\text{ENG},0,3 \text{ the plant}} \right] \left[ \text{before}_F \left[ \text{DP}_1 \text{ the time} \left[ \text{which} \lambda_5 \lambda_3 \left[ \text{Sally say-past}^{\text{ENG},0,2 \text{ it bloom-past}^{\text{ENG},6,7 \text{ it bloomed}}} \right] \right] \right] \right] \)

b. \([\text{at} e_5]^{C,g}@,g\) modifies \( S1 \) (= \( [S1 \ldots \text{Sally say} \ldots [S2 \ldots \text{it bloom} \ldots ]] \))

c. \([\text{[DP}_1]\]^{C,g}@,g[0-SP]\) is the unique \( C \)-relevant \( t \in D_i \) such that there is a \( t' \in D_i \) such that \( t' < \text{SP} \) and Sally says “the plant bloomed” in \( @ \) at \( t' \) and \( t' = t \).

d. The watering time in \( @ \) (which precedes SP) precedes \([\text{[DP}_1]\]^{C,g}@,g[0-SP]\).

(93) Low reading with \([\text{before}_F]^{C,g}\) – NPI blocked by the time (see (88))

a. \( \exists \left[ \lambda_3 \left[ \text{John water-past}^{\text{ENG},0,3 \text{ the plant}} \right] \left[ \text{before}_F \left[ \text{[DP}_2 \text{ the time} \left[ \text{which} \lambda_5 \lambda_3 \left[ \text{Sally say-past}^{\text{ENG},0,2 \text{ it bloom-past}^{\text{ENG},6,7 \text{ it bloomed}}} \right] \right] \right] \right] \)

b. \([\text{at} e_5]^{C,g}@,g\) modifies \( S2 \) (= \( [S2 \ldots \text{it bloom} \ldots ] \))

c. \([\text{[DP}_2]\]^{C,g}@,g[0-SP]\) is the unique \( C \)-relevant \( f \in D_{i<s,i>} \) such that for all worlds \( w \) and times \( t \) compatible with what Sally says in \( @ \) at \( g(2) \) (< SP), there is a time \( t' < t \) such that the plant bloomed in \( w \) at \( t' \) and \( t' = f(t)(w) \).

d. When \([\text{[DP}_2]\]^{C,g}@,g[0-SP] = “1pm”, 1pm is the time the plant bloomed according to Sally; when \([\text{[DP}_2]\]^{C,g}@,g[0-SP] = “Grandma’s death time”, Grandma’s death time according to Sally is also the time the plant bloomed according to Sally.

e. The watering time in \( @ \) (which precedes SP) precedes \([\text{[DP}_2]\]^{C,g}@,g[0-SP])(SP)(@).

15 \([\text{say}]^{C,g}@,g\) and \([\text{before}_B^{C,g}]^{C,g}@,g\) always apply to the intension of their complement, by Heim and Kratzer’s (1998) Intensional Functional Application (IFA, see (i)). As for \( \text{at} \) and \( \text{before}_F \), in (92), \([\text{at}]^{C,g}@,g\) takes \([e_5]^{C,g}@,g\) (for all relevant \( t \)) and \([\text{before}_F]^{C,g}@,g\) takes \([\text{[DP}_1]\]^{C,g}@,g[0-SP]\). By rule (ii), in (93) \([\text{at}]^{C,g}@,g(\text{at} e_5]^{C,g}@,g(\text{at} e_5]^{C,g}@,g(\text{at} e_5]^{C,g}@,g(\text{at} E_5]^{C,g}@,g(\text{at} E_5]^{C,g}@,g(\text{at} E_5]^{C,g}@,g\text{times} t \) and worlds \( w \) and \([\text{before}_F]^{C,g}@,g\) takes \([\text{[DP}_2]\]^{C,g}@,g[0-SP]@)\).
Crucially, there is no way to obtain any low reading – ‘de re’ or ‘de dicto’ – from (91a), because it does not have any pronominal element of type <i, <s, i>> that can be abstracted over.

A prediction regarding intensionality: \(\text{before}^F\), unlike \(\text{before}^{B&C}\), is by assumption strictly “extensional”, in the sense that its semantics (see (90)) makes no reference to alternative worlds. Indeed, (94) forces \(p\)-factivity, entailing that the plant died (unlike \text{John watered the plant before it died}, which does not; see discussion of (13a) in Section 2).

(94) John watered the plant before the time at which the plant died.

This is also reflected in the Spanish data above: In (82), \text{think} appears with subjunctive morphology (in keeping with the intensionality of \(\text{before}^{B&C}\)); in (83), \text{think} appears with indicative morphology (in keeping with the non-intensionality of \(\text{before}^F\)).

Some additional cross-linguistic consequences. It is observed in Otaki (2010) that the acceptability of Japanese past-under-past improves dramatically in \(\text{before}\)-clauses with an overt \text{time at which}, as in (95a), which is not perfect, but contrasts sharply with its ungrammatical “bare” \(\text{before}\) counterpart. If mae-ni in this example is indeed \(\text{before}^F\) (see (95b)), this is expected on the theory advocated here, because \(\text{before}^F\) lacks \text{EARLIEST}, so \(\text{IPF}\) is avoided. (96a), with present-under-past also has an OK-but-less-than-perfect status (Koichi Otaki, p.c.), indicating that in general “bare” \(\text{before}\)-clauses are preferred.

(95) a. Masa-wa [Ken-ga ki-ta jikan go-hun mae-ni] kaet-ta
   Masa-TOP Ken-NOM come-PAST time five.minutes before-at return-PAST
   ‘Masa returned five minutes before a time which Ken came at.’
   b. \([t_0 \text{[PAST}^{asp} [[\lambda_1 [Masa \text{return}-t_1]] \text{[before}^F \text{[the time [which } \lambda_5 \text{[t}_5 \text{[PAST}^{asp} [[\lambda_3 [Ken \text{come-t}_3]] \text{[at } e_5]]]]]]]]]]]]\)

(96) a. Masa-wa [Ken-ga ku-ru jikan go-hun mae-ni] kaet-ta
   Masa-TOP Ken-NOM come-PRES time five.minutes before-at return-PAST
   b. \([t_0 \text{[PAST}^{asp} [[\lambda_1 [Masa \text{return}-t_1]] \text{[before}^F \text{[the time [which } \lambda_5 \text{[t}_5 \text{[PRES}^{asp} [[\lambda_3 [Ken \text{come-t}_3]] \text{[at } e_5]]]]]]]]\]

As expected, (95a) obeys No-\(p\)-shiftability, as the evaluation time of the embedded past is \(SP\), so it entails that Ken comes prior to \(SP\). (96a), where the evaluation time of the embedded present is bound, exhibits \(p\)-shiftability just like its “bare” \(\text{before}\) counterpart (despite \(p\)-factivity, enforced by \(\text{before}^F\)), and may be used to mean that Ken will come in the future or to imply that the speaker is agnostic about the exact time of Ken’s coming (see discussion of (51) in 3.3).

What still remains a bit mysterious is the fact that the degree phrase go-hun is required; without it both (95a) and (96a) are again ungrammatical.
(see Takahashi 2007 for related observations). However, as also noted in Otaki (2010), the presence of a degree phrase supports Geis-ambiguities (even without an overt the time at which), as shown by the contrast in (97).

\[(97)\] 
\[\text{Watasi-wa Satoshi-ni [anata-ga [kaet-ta to] it-ta ni-jikan mae-ni]}\]
\[\text{Amherst-de at-ta. Amherst-in meet-PAST} \]
\[\text{‘I met Satoshi in Amherst two hours before you said (he) left.’ HIGH/LOW}\]

\[\text{Watasi-wa Satoshi-ni [anata-ga [kaet-ta to] i-u mae-ni]}\]
\[\text{Amherst-de at-ta. Amherst-in meet-PAST} \]
\[\text{‘I met Satoshi in Amherst before you said (he) left.’ HIGH/\ast LOW}\]

This suggests that Japanese does sometimes allow a covert time at which, but for some (yet unknown) reason, both overt and covert time at which require the presence of a degree phrase (see Otaki 2010 for discussion). Importantly, it is hard to see how the English/Japanese contrast regarding past-under-past under before can be accounted for without appealing to an inherent difference in the tenses themselves. For example, Grønn & von Stechow (to appear), who do not acknowledge that both English and Japanese have two kinds of before-constructions, adopt an LF along the lines of the relative clause LF in (80a) for English before-constructions (all of which, according to them, contain the time at which overtly or covertly), and an LF along the lines of (36a) for Japanese (which, according to them, does not allow the time at which to precede before). They account for the unacceptability of part-under-past in Japanese before-clauses (without the time at which) by claiming that (36a) leads to a pragmatic anomaly. Indeed, as we saw in Section 3.1 (see discussion of (27)), if we take C to be a set of intervals (as Grønn & von Stechow would), rather than an interval, we derive a meaning for ‘q before-C p’ that does not guarantee q-precedence. According to Grønn & von Stechow, the resulting meaning is misleading because it makes the contribution of before uninformative. However, once we acknowledge that both languages can, in principle, have before-clauses with and without the time at which, and given that the English NPIs facts suggest that before-clauses with NPIs do not contain the time at which, we expect – on the Grønn & von Stechow view – that English past-under-past before-clauses with NPIs be disallowed (on a par with Japanese), but this is not so (a similar problem arises for the proposal in Krifka 2010, which does not rely on EARLIEST).
We do not attempt to capture any additional generalizations about Geis-ambiguities (leaving this for further research), and apart from this section, all our predictions concern before-clauses without the time at which (overt or covert). We now return to these.

6 MORE ON CROSS-LINGUISTIC VARIATION

6.1 Permissible and impermissible languages

What the theory outlined in Sections 3 and 4 amounts to, in effect, is that all languages have the same before (namely, before^{b&c}) and the same say/believe, but they may differ along the following four dimensions: (a) having an SOT rule; (b) having quantificational tenses; (c) having pronominal tenses; and (d) having a “shiftable” present tense. There are many conceivable ways in which these dimensions could interact with each other, which means that there are many conceivable ways in which language variation may be restricted. Let us elaborate on this point.

The first question that comes to mind is whether languages can have no tenses at all. This is a very difficult question, as it is well known that some languages indeed do not have overt morphological tenses, and it is far from clear that such languages still have covert tenses. We do not try to settle this issue here (see Matthewson 2006 and Bittner 2011 for two opposing views), but make the more modest suggestion that if a language has tenses, they have to be pronominal or quantificational, but they cannot be neither pronominal nor quantificational.

Secondly, we assume (some version of) the Embeddability Principle, in the sense of Sharvit (2003, 2008). It is argued there that there cannot be a language where it is impossible to report the following state of affairs via embedding: Two hundred years ago John had the following thought: “seven is not a prime number.” In English, such a state of affairs is reported by John thought that seven was not a prime number, by the application of the SOT rule to the embedded past. In languages that have a shiftable present tense, such an event may be reported with present-under-past (see Sharvit 2008 and Ogihara & Sharvit 2012 for discussion of why applying the ‘de re’ strategy to an embedded past is inapplicable in this case; cf. (66a)). Without the Embeddability Principle, we would predict there to be languages that have no SOT rule and no shiftable present. That is to say, without the Embeddability Principle we would predict there to be languages where the only way to report such a state of affairs is via direct quotation; the Embeddability Principle prohibits such languages. We assume that some version of the Embeddability Principle is also applicable to before-clauses, in the
following sense: \([S \text{ before}^{B&C} S']\) (where \(S\) and \(S'\) stand for sentence LF s “referring” to two non-overlapping events) is well-formed in every language that has \(\text{before}^{B&C}\). That is to say, a language that has \(\text{before}^{B&C}\), only quantificational tenses, no SOT rule and no shiftable present is not a possible language, because in such a language a state of affairs where John left, Mary arrived, and the former event preceded the latter can only be reported with “John left, Mary arrived, and the first event preceded the second” or some such similar paraphrase. It is our contention that such natural languages do not exist. In other words, we assume that just like every language must have the syntactic tools for embedding under an attitude verb, it must also have the syntactic tools for embedding under \(\text{before}^{B&C}\) (if it has this lexical item at all).

Suppose we also assumed (perhaps due to some Economy principle) that a language cannot have both pronominal and quantification tenses. This would still leave some room for variation, as shown in (98), which permits six types of languages-with-tenses, with English being type 6, Polish type 10, and Japanese type 11. Clearly, if we assume that languages can have pronominal and quantificational tenses we get nine permissible combinations (as types 1, 5, and 9, which are impermissible according to (98), become permissible).

\[
\begin{array}{cccc}
\text{Type} & \text{SOT-rule} & \text{Shiftable present} & \text{Pronominal tense} & \text{Quantificational tense} \\
1. & * & + & + & + \\
2. & \checkmark & + & + & -- \\
3. & \checkmark & + & -- & + \\
4. & * & + & -- & -- \\
5. & * & -- & + & -- \\
6. & \checkmark & -- & + & -- \\
7. & \checkmark & -- & -- & + \\
8. & * & -- & -- & -- \\
9. & * & -- & + & + \\
10. & \checkmark & -- & + & -- \\
11. & \checkmark & -- & -- & + \\
12. & * & -- & -- & -- \\
13. & * & -- & -- & + \\
14. & * & -- & -- & + \\
15. & * & -- & -- & + \\
16. & * & -- & -- & -- \\
\end{array}
\]

(98) does not distinguish between full shiftability and semi-shiftability. We suggested in 4.2 Japanese has quantificational tenses with a fully shiftable present, and Polish has pronominal tenses with a semi-shiftable present. The theory does not ban languages with a fully shiftable pronominal present. On the other hand, a quantificational present which is “semi-shiftable” is not possible, as we saw in connection with (77). If Japanese had such a “semi-shiftable” present it would violate the Embeddability Principle. Notice that we will get a picture that is
somewhat different from the one in (98) if we assume that languages can have mixed systems (for example, a quantificational past and a pronominal present, or vice versa); such an option is missing from (98) altogether (and might provide another way of complying with the Embeddability Principle).  

We do not have examples of all the languages that are predicted to be permissible (even by (98), which is relatively strict), but it is not unreasonable that accidental gaps occur. We also do not attempt to decide between the possible permissible parametric combinations. Rather, we would like this to be thought of as a research program, with the hope that once additional cross-linguistic data become available, the choice between the conceivable parametric combinations will become clearer. Importantly, any variant of our theory – even the one in (98) – makes clear predictions regarding (No-)p-shiftability in \textit{before}-clauses, as summarized in (99). (99a,b,c) are a direct result of our assumptions regarding the meaning of \textit{before} and how it interacts with tense parameters.

\begin{equation}
\begin{aligned}
\text{(99) a. If a language L has a well-formed present-under-past in } \textit{before}-
\text{clauses, then L exhibits p-shiftability, rather than No-p-shiftability, in present-under-past } \textit{before}-
\text{clauses.} \\
\text{(Why? A well-formed present-under-past in } \textit{before}-
\text{clauses implies that the present is fully shiftable.)} \\
\text{b. If a language L has a well-formed past-under-past in } \textit{before}-
\text{clauses and it exhibits p-shiftability, rather than No-p-shiftability, then L has “simultaneous” readings of past-under-past in attitude reports.} \\
\text{(Why? A past-under-past in } \textit{before}-
\text{clauses with p-shiftability, implies a “null” interpretation of the past tense, which is achieved via the SOT rule.)} \\
\text{c. If a language L has a well-formed past-under-past in } \textit{before}-
\text{clauses and no “simultaneous” readings of past-under-past in attitude reports, then L exhibits No-p-shiftability, rather than p-shiftability, in past-under-past } \textit{before}-
\text{clauses.} \\
\text{(Why? No “simultaneous” readings implies no SOT-rule, which implies that the past-under-past in } \textit{before}-
\text{clauses is pronominal; as such it is not “null” in } \textit{before}-
\text{clauses and leads to No-p-shiftability.)}
\end{aligned}
\end{equation}

\footnote{An anonymous reviewer wonders whether languages may differ with respect to the shiftability of the past tense. We assume the past is universally shiftable (simply because we have had no empirical reason to assume otherwise), but this it is certainly a logical possibility that languages differ with respect to this as well.}
Japanese illustrates (99a), and Polish illustrates (99c). It seems that some dialects of Spanish illustrate (99b). For speakers of those dialects, (100) is fine when Maria hasn’t arrived yet, and (101) (which means ‘Maria will leave only tomorrow. Juan arrived before Maria’s leaving’) is coherent: Maria’s leaving can occur in the future.

(100) Juan se fue antes de que llegara María
     Juan CL go-PAST before of that arrive-SUBJ-PAST Maria
(101) María recién se va a ir mañana.
     Maria just CL going-to-go tomorrow.
     Juan llegó antes de que se fuera (María)
     Juan arrive-PAST before of that CL leave-PAST-SUBJ Maria

For other speakers of Spanish, (100) is acceptable only if Maria’s arrival precedes SP, and (101) is incoherent. Let us call the dialect that exhibits No-p-shiftability Spanish A and the one that exhibits p-shiftability Spanish B. Since both dialects have the SOT-rule (see, for example, Gennari 1999; see also 6.2), it seems that the contrast between them lies in the nature of the matrix past tense in (101): in Spanish B the past tense is quantificational, so it QRs and triggers “deletion” of the past tense in the before-clause; in Spanish A the past tense is pronominal, so it cannot trigger “deletion” of the past tense in the before-clause. This means that in addition to predicting more than one type of non-SOT languages (Japanese and Polish), we also predict more than one type of SOT languages: Spanish A is an SOT language with acceptable past-under-past before-clauses with No-p-shiftability; Spanish B is SOT with acceptable past-under-past before-clauses with p-shiftability.

6.2 The limitations of the proposal

As it turns out (and as the readers may have figured out), our theory does not suffice to predict all permissible varieties (even of languages with overt tenses). In particular, our theory seems to predict the following language type to be impermissible.

(102) √ “null” past-under-past in attitude reports; *past-under-past in before-clauses

The reason is this: a “null” past-under-past in attitude reports means that the language has the SOT rule. If its past tense is quantificational (e.g., #7 (98)), then it should be OK in before-clauses as well (triggering

17 Among the speakers who were consulted, all River Plate Spanish speakers allowed p-shiftability, and all Peninsular Spanish speakers enforced No-p-shiftability. However, it is not obvious that the “official” River Plate/Peninsular divide indeed corresponds to the Spanish A/Spanish B divide, as we describe it here.
“deletion” of the embedded past); if its past tense is pronominal (e.g., #6 in (98)), then it should be licensed in before-clauses (just like it is in English and Polish). It seems that Modern Greek is a (102)-language. (103) shows that past-under-past in before-clauses is ungrammatical, and (104), with a past-under-past attitude report, has a “simultaneous” reading (showing that past-under-past in attitude reports can be “null”).

(103) *O Janis efige prin (na) efige/efevge i Maria
the John leave-PAST-PERF before SUBJ leave-PAST-PERF/IMPERF the Mary

(104) To 1963 o Kostas mas ipe oti i Maria itan eggios
the 1963 the Kostas us tell-PAST that the Maria be-PAST pregnant

The challenge becomes even bigger when we consider the fact that Spanish dialects A and B both have past-under-past, with subjunctive marking, as we saw in 6.1. This means that there is no inherent incompatibility between past subjunctive and before. Clearly, what is missing from (98) is some indication as to how tense parameters interact with parameters pertaining to mood distinctions (subjunctive vs. indicative), and/or parameters pertaining to aspectual distinctions (perfective vs. imperfective). There are many conceivable forms that these interactions could take; let us sketch one of them in rough terms (only to show that the existence of Modern Greek doesn’t necessarily force us to abandon our proposal). Our rough and sketchy discussion relies on the assumption that there is an independent theory of the subjunctive, which accounts for why certain languages have subjunctive marking and others do not, and for why, in each language, this marking appears only in some intensional constructions but not others.

Suppose subjunctive mood comes (universally) in two varieties: semantic and morphological. In languages with pronominal tenses semantic subjunctive results in obligatory binding of the evaluation index of a pronominal tense (obviating any indexing constraints, if the language has any). Morphological subjunctive has no such effect, it is simply a process that affects verbal morphology in some intensional environments. Suppose further that a language can have a pronominal present and a quantificational past (or the other way around), and that Modern Greek, Spanish A and Spanish B (see 4.2) have the following parameter settings (for current purposes, let us assume that Modern Greek ‘before’-clauses have an underlying subjunctive even when it is not pronounced).

(105) a. Modern Greek: +SOT; fully/semi-shiftable pronominal present; pronominal past; semantic subjunctive

b. Spanish A: +SOT; non-shiftable pronominal present; pronominal past; morphological subjunctive
c. Spanish B: +SOT; non-shiftable pronominal present; quantificational past, semantic subjunctive

The predictions are therefore as follows: present-under-past is excluded in Spanish A before-clauses because given that it has only morphological subjunctive, there is no way to circumvent the ban on identical indices. Past-under-past is excluded in Modern Greek before-clauses because binding of the evaluation-index of the pronominal past results in a presupposition failure (cf. (42)). In Spanish B, on the other hand, which has semantic subjunctive and a quantificational past, both past- and present-under-past are possible in before-clauses.

(106) Spanish A – morphological SUBJ
   a. √ ... past^{SPN-A}_{0, j} ... before [λ_j [... M-SUBJ]-past^{SPN-A}_{0, j} ....
      No effect on the embedded indices.
   b. * ... past^{SPN-A}_{0, j} ... before [λ_j [... M-SUBJ]-pres^{SPN-A}_{0, j} ....
      Stump effect

Modern Greek – semantic SUBJ
   a. * ... past^{GREK}_{0, j} ... before [λ_0 [... S-SUBJ]-past^{GREK}_{0, j} ....
      Semantic SUBJ binds ‘0’, resulting in a presupposition failure; cf. (42)
   b. √ ... past^{GREK}_{0, j} ... before [λ_j [... S-SUBJ]-pres^{GREK}_{j, j} ....
      Semantic SUBJ binds ‘j’: No presupposition failure; cf. (44)

Spanish B – semantic SUBJ
   a. √ [t_0 [PAST^{SPN-B} ... [before [λ_0 [... S-SUBJ]-PAST^{SPN-B} ...
      Semantic SUBJ binds ‘j’: No presupposition failure; cf. (44)
   b. √ [t_0 [PAST^{SPN-B} ... [before [λ_0 [... S-SUBJ]-pres^{SPN-B}_{j, j} ....

Crucially, Modern Greek and Spanish B are both predicted, by (105), to allow present-under-past in before-clauses: both have semantic subjunctive. For Modern Greek, this is illustrated in (107); for Spanish B, this is evidenced by the fact that (108) is grammatical in Spanish B and ungrammatical in Spanish A.

(107) O Janis efige prin (na) figi/*fevgi i Maria
    the John leave-PAST-PERF before SUBJ leave-PRES-PERF/IMPERF the Mary

(108) Juan se fue antes de que llegue María
    John CL go-PAST before of that arrive-SUBJ-PRES Mary

As we saw in 4.1, in connection with (100) and (101), Spanish A and Spanish B contrast regarding (No-)p-shiftability of the past tense: Spanish B exhibits p-shiftability and Spanish A No-p-shiftability. This is compatible with the assumptions in (106), as is the fact that the
embedded present in Modern Greek and Spanish B exhibit p-shiftability, as evidenced by the fact that Mary’s arrival in (107) and (108) need not precede SP.

That Modern Greek has both the SOT rule and a shiftable present tense is consistent with what is documented in, for example, Schlenker 1999, and supported by the fact that both past-under-past and present-under-past support “simultaneous” readings of attitude reports, as shown in (109): both variants have the “simultaneous” reading. That both dialects of Spanish have the SOT rule but no shiftable present tense is shown by the fact that (110a) has a “simultaneous” reading but (110b) only a “double access” reading. Finally, (105) is also consistent with the fact that the present tense in Modern Greek and in (both dialects of) Spanish obligatorily refers to SP in present-under-past relative clauses, as shown by (111) (where the travel-loving time overlaps SP and by the oddity of (112) (alfajor is a kind of cream-filled cookie).

(109) a. To 1963 o Kostas mas ipe oti i Maria ine eggios
   the 1963 the Kostas us tell-PAST that the Maria be-PRES pregnant
   b. To 1963 o Kostas mas ipe oti i Maria itan eggios
   the 1963 the Kostas us tell-PAST that the Maria be-PAST pregnant

(110) a. Juan dijo que Maria estaba enferma
   Juan say-PAST-PRF that Maria be-PAST-IMP sick
   b. Juan dijo que Maria está enferma
   Juan say-PAST-PRF that Maria be-PRES sick

(111) stin pediki tu ilikia, o Janis gnorise mia gineka pu agapai na taksidevi
   in.the child his age, the John meet-PAST a woman that love-PRES to travel
   ‘In his childhood, John met a woman who loves traveling’

(112) #Juan se comió todo el alfajor que está sobre la mesa.
   Juan CL eat-PAST all the alfajor that be-PRES on the table
   ‘Juan ate all the alfajor that is on the table’

There are probably many more parameters whose interaction with tense parameters could, in principle, affect the before-pattern. Our purpose in this section has merely been to illustrate this point and show how the problem could be approached. A more general theory would require a careful examination of each of these parameters.

Finally, the discussion of the Modern Greek and Spanish data raises some thoughts about cases of well-formed English present-under-past before-clauses, which seem to be inconsistent with the unacceptability of (2). Apparently, in a context where the plant has a good chance of dying in the near future, one can felicitously utter (113) with an embedded present.

(113) #Juan se comió todo el alfajor que está sobre la mesa.
   Juan CL eat-PAST all the alfajor that be-PRES on the table
   ‘Juan ate all the alfajor that is on the table’
John (just) watered the plant before it dies/died (on him).

It is plausible that this occurrence of the present tense comes accompanied by a covert subjunctive morpheme, which undoes the indexing constraint (the same constraint that excludes (44)), as we hypothesized for Greek and Spanish B. If this analysis is on the right track, it would have to rely on a theory of the subjunctive that predicts, independently, that the subjunctive in English before-clauses is not obligatory, and that on their own, say and ask cannot license the subjunctive; otherwise John said Mary is pregnant would have more than just the “double access” reading (see 3.2).

7 SUMMARY

It is argued in Ogihara & Sharvit (2012) that the fact that non-SOT languages come in more than one variety can be explained by assuming that languages can differ with respect to the type of tense they have. Here we have re-interpreted this idea in order to account for cross-linguistic variation in before-clauses, distinguishing between non-SOT languages that have an English-like pattern in before-clauses, and non-SOT languages that do not. It was shown that it is the earliest-operator, which is presumably involved in the semantics of before, that enables us to exploit the quantificational/pronominal distinction in order the predict the observed variation: (i) a quantificational “non-deleted” past leads to IPF, and IPF is avoided either by the availability of a pronominal tense or by the availability of the SOT rule; (ii) an absolute present results in the Stump effect; a shiftable present need not.

An important ingredient of the theory is the SOT-rule which, as we saw, is applicable in some constructions but not in others (as it requires that the “deleting” tense c-command the “deleted” tense). For example, it is not applicable in English before-clauses, unless the entire sentence – ‘q before p’ – is further embedded by an attitude verb. It is hard to see how any theory that does not acknowledge the SOT rule (e.g., Kubota et al. 2011), can account for the fact that English before-clauses “acquire” p-shiftability when ‘q before p’ is embedded under an attitude verb (see (66b) and (67) in Section 4).

Let us end with an additional unsolved puzzle. Do other temporal clauses – specifically, when-clauses and after-clauses – involve the earliest-operator (or a similar operator)? The simplest and most elegant theory would say that they do, predicting straightforwardly the before-pattern to be exhibited by all temporal clauses. It turns out that this prediction is incorrect. Here we simply point out that this suggests that
not all temporal clauses involve an \textit{EARLIEST}-like operator. A detailed analysis of other temporal operators is beyond the scope of this paper, but it is worth noting that Japanese \textit{when}-clauses and \textit{after}-clauses with \textit{past-under-past} (see Ogihara 1996 and Arregui & Kusumoto 1998 among others), unlike \textit{past-under-past} \textit{before}-clauses, are grammatical. This would be unexpected on the assumption that the semantics of \textit{when} involves an operator that produces the same semantic anomaly for embedded quantificational past tenses as the \textit{EARLIEST}-operator. But on the assumption that the semantics of \textit{when} and \textit{after} involve a different operator, this state of affairs could be theoretically unproblematic.

\section*{Acknowledgments}

I am immensely grateful to Toshiyuki Ogihara and Lyn Shan Tieu, whose co-authored works with me – Ogihara & Sharvit (2012) and Sharvit & Tieu (2011) respectively – inspired this paper. I am equally indebted to Koichi Otaki for sharing with me his work and his insights on Japanese tense, and to Gidi Avrahami for his invaluable feedback on the formal analysis. For comments, questions and native speaker judgments, I thank the audiences at UCLA, UCSD and Rome Science Festival 2012, as well as the following people: Daniel Altshuler, Ana Arregui, Ivano Caponigro, Barbara Citko, Sam Cumming, Veneeta Dayal, Ben George, Bart Geurts, Gwen Gillingham, Mary Goodrich, Corina Goodwin, Mira Goral, Ate Grønn, Elena Guerzoni, Dita Gutman, Bruce Hayes, I-Ta Hsieh, Sabine Iatridou, Makoto Kanazawa, Jungmin Kang, Jasmeen Kanwal, Ed Keenan, Nathan Klinedinst, Natasha Korotkova, Helen Koulidobrova, Diego Krivochen, Roger Levy, Anna Lubowicz, Peter Ludlow, Daniel Margulies, Ned Markosian, Luisa Martí Martínez, Kimiko Nakanishi, Florencio Noceti, Toshiko Oda, Jessica Rett, Jacopo Romoli, Daniel Rothschild, Ariel Rubinstein, Tsuyoshi Sawada, Carson Schütze, Scott Seyfarth, Yoshiyuki Sibata, Benjamin Spector, Ed Stabler, Jason Stanley, Masahiko Takahashi, Jos Tellings, Mookie Tenembaum, Arhonto Terzi, Guillermo Toscano y García, Guillaume Thomas, Julio Villa-García, Kaeli Ward, Lauren Winans, Ting Xu, Sandro Zucchi and five (!) anonymous \textit{Journal of Semantics} reviewers. I alone am responsible for any and all typos, errors, inaccuracies, misunderstandings or misconceptions.

\textbf{YAEL SHARVIT}

\textit{Department of Linguistics}  \\
\textit{3125 Campbell Hall}  \\
\textit{University of California, Los Angeles}  \\
\textit{Los Angeles, CA 90095-1543, USA}  \\
e-mail: ysharvit@gmail.com
APPENDIX

One of the semantic properties of sentences of the form ‘q before p’ discussed in Section 2 is No-p-factivity, illustrated by the fact that John watered the plant before it died does not entail that the plant dies in @ (the actual world). A concrete proposal for the intensional meaning of before might be (1), where p is an element of D_{<s, <i, t>}, t an element of D_{i}, W is the set of all possible worlds, and for any w ∈ W and t ∈ D_i, ACC_w,t is the set of worlds accessible from w at t.

(1) \[[before]_{B&C-INS}^{s, i, t} \] C,w,g(p)(t) is defined only if:
(a) t ⊆ C and ACC_w,t ≠ ∅;
(b) ACC_w,t ⊆ {w' ∈ W | (i) EARLIEST_C({t' ∈ D_{i} | p(w')(t') = True}) is defined; and
(ii) MIN(C) < EARLIEST_C({t' ∈ D_{i} | p(w')(t') = True})}; and
(c) if {t' ∈ D_{i} | p(w)(t') = True} ≠ ∅, then w ∈ ACC_w,t.
When defined, \[[before]_{B&C-INS}^{s, i, t} \] C,w,g(p)(t) := True iff ACC_w,t ⊆ {w' ∈ W | t < EARLIEST_C({t' ∈ D_{i} | p(w')(t') = True})}.

In the course of the semantic computation of the LF \[∃ \[[\lambda_3 \[[John]_{ENG}^{0,3} \text{water-past}] \] \[before]_{B&C-INS}^{s, i, t} \] \[\lambda_2 \[[it]_{ENG}^{0,2} \text{die-past}] \] \] (the LF of John watered the plant before it died), Intensional Functional Application (as formulated, for example, in Heim & Kratzer 1998) is applied to before_{B&C-INS} and \[[\lambda_2 \[[it]_{ENG}^{0,2} \text{die-past}] \] \], which is of type <s, <i, t>-, guaranteeing that before_{B&C-INS} takes a <s, <i, t>,-argument (the function that maps every world w to \[\[[\lambda_2 \[[it]_{ENG}^{0,2} \text{die-past}] \] \] w,C,g[0->SP]). Thus, when John watered the plant before it died is evaluated relative to @, it is possible for ‘the plant dies’ to be false in @, at any past time (though it cannot be true in @ prior to when it is true in the worlds accessible from @).

However, (1) does not account for before-data with degree phrases and specifically, the interaction between the intensionality of before and the semantic contribution of degree phrases. As observed in Heinämäki (1974), the presence of a degree phrase biases a sentence of the form ‘q deg-before p’ towards p-factivity. This is supported by the fact that (2b) is odd (the degree phrase forces a factive reading which, based on real-world knowledge, describes an impossible state of affairs), as opposed to the ambiguous (2a).

(2) a. Clint died before he met his grandchildren.
   Factive reading: Clint actually met his grandchildren.
   Non-factive reading: Clint never met/meets his grandchildren.

b. #Clint died two weeks before he met his grandchildren.

To account for these facts, we first distinguish between elements of D_{i} and their degrees of length – elements of D_{v}. Accordingly, two
distinct intervals may have the same length (for example, \([3, 4] \neq [4, 5]\), but \(\text{LENGTH}([3, 4]) = \text{LENGTH}([4, 5]) = 1\)). Secondly, we suggest that the meaning of \(\text{before}\) is as in (3), according to which it takes an additional argument, \(f\), an element of \(D_{s, v}\) (a function from worlds to degrees). (3) refers to ‘\(\oplus\)’, defined in (4).

(3) \(\text{before}_{B \& C-\text{DEG}}^{w, C, g}(p)(f)(t)\) is defined only if:
(a) \(t \subseteq C\) and \(\text{ACC}_{w,t} \neq \emptyset\);
(b) \(\text{ACC}_{w,t} \subseteq \{w' \in W|\)
   (i) \(\text{EARLIEST}_C(\{t' \in D_i| p(w')(t') = \text{True}\})\) is defined;
   (ii) \(f(w')\) is larger than zero;
   (iii) \(\text{MIN}(C) < \text{EARLIEST}_C(\{t' \in D_i| p(w')(t') = \text{True}\})\);
   (c) if \(\{t' \in D_i| p(w)(t') = \text{True}\} \neq \emptyset\), then \(w \in \text{ACC}_{w,t}\); and
   (d) if there is a \(d \in D_v\) larger than zero such that \(\text{ACC}_{w,t} \subseteq \{w' \in W| f(w') \geq d\}\), \(w \in \text{ACC}_{w,t}\).
When defined, \(\text{before}_{B \& C-\text{DEG}}^{w, C, g}(p)(f)(t) := \text{True} \iff \text{ACC}_{w,t} \subseteq \{w' \in W| \text{MAX}(\text{MIN}(t) \oplus f(w')) \leq \text{EARLIEST}_C(\{t' \in D_i| p(w')(t') = \text{True}\})\}\).

(4) For any \(m \in M\) and \(d \in D_v\), \(m \oplus d\) is the unique \(t \in D_i\) such that:
(i) \(\text{MIN}(t) = m\); and (ii) \(\text{LENGTH}(t) = d\).

As (3) implies, \(\text{before}\) always takes an \(f\)-argument, even when that argument is not phonetically realized. Accordingly, the following two LFs are generated for (2a,b) – the LF in (5) and the LF in (6). The meaning of \(\text{two weeks}_j\) is the Heimian meaning (Heim 1988a) in (7a); \(\text{DEG}_7\) in (6) is a silent pronoun, with the meaning in (7b).

(5) \(\exists \left[ [\lambda_3 \ [\text{Clint die-past}^{\text{ENG}, 0, 3}][[\text{DEG} \text{two weeks}_5]] \ [\text{before}_{B \& C-\text{DEG}}^{w, C, g}] [\lambda_2 \ [\text{he meet-past}^{\text{ENG}, 0, 2} \text{his grandchildren}]]] \right]\)

(6) \(\exists \left[ [\lambda_3 \ [\text{Clint die-past}^{\text{ENG}, 0, 3}][[\text{DEG} \text{DEG}_7]] \ [\text{before}_{B \& C-\text{DEG}}^{w, C, g}] [\lambda_2 \ [\text{he meet-past}^{\text{ENG}, 0, 2} \text{his grandchildren}]]] \right]\)

(7a) \(\left[ [\text{two weeks}_j]\right]^{C, w, g} \text{ is defined only if } g(j) \in D_v \text{ and } g(j) \text{ is at least “two weeks”}.\)
   When defined, \(\left[ [\text{two weeks}_j]\right]^{C, w, g} = g(j)\).

(7b) \(\left[ [\text{DEG}_j]\right]^{C, w, g} \text{ is defined only if } g(j) \in D_{<s, v>}\).
   When defined, \(\left[ [\text{DEG}_j]\right]^{C, w, g} = g(j)\).

In (5), \(\text{before}_{B \& C-\text{DEG}}^{w, C, g} [\lambda_2 \ [\text{he meet-past}^{\text{ENG}, 0, 2} \text{his grandchildren}]]\) combines with \(\text{two weeks}\) by Intensional Functional Application: its denotation applies to the function that maps every world \(w\) to \(\left[ [\text{two weeks}_5]\right]^{w, C, g}\). In (6), \(\text{before}_{B \& C-\text{DEG}}^{w, C, g} [\lambda_2 \ [\text{he meet-past}^{\text{ENG}, 0, 2} \text{his grandchildren}]]\) combines directly with \(\text{DEG}_7\) by Functional Application.
When evaluated relative to @, (5) entails that for all \( w' \in \text{ACC} \), two weeks \( \leq (w') \) is at least two weeks, so it follows, by (3d), that \( @ \in \text{ACC} \), leading to the implication that Clint actually met his grandchildren. Following Landman (2004), we assume that two days and exactly two days have the same meaning, but the presence of exactly induces a non-defeasible ‘exactly’–implicature. For example, John watered the plant exactly two days before it died comes with the non-defeasible implicature according to which all stronger alternatives to John watered the plant two days before it died (namely, John watered the plant two and a half days before it died, John watered the plant three days before it died, etc.) are false.

As for (6), there are contexts/assignments where zero is the largest \( d \in \text{D_v} \) such that for all \( w' \in \text{ACC} \), \( \text{DEG} \leq (w') \geq d \), so it does not follow that \( @ \in \text{ACC} \). This, in turn, implies that according to (6), although Clint meets his grandchildren in the worlds accessible from @, @ itself need not be among these worlds (furthermore, there is no requirement that ‘Clint dies’ be true in the worlds accessible from @). This prediction is consistent with the fact that not all degree phrases force p–factivity, as illustrated by Clint died just before he met his grandchildren, which does not entail that Clint actually met his grandchildren (Ed Keenan, p.c.). Whatever the exact meaning of just is, unlike two weeks it does not “refer” to a specific minimal degree.

REFERENCES


Bäuerle, Rainer. (1977), Tempus, Temporaladverb und die Temporale Frage. Papers 13 and 15 of SFB, University of Konstanz, Germany.


Bennett, Michael & Barbara Partee. (1972), Toward the Logic of Tense and Aspect in English. Ms. [Published 1978. Indiana University Linguistics Club, Bloomington, Indiana.]


von Stechow, Arnim. (2009), ‘Tenses in compositional semantics’. In W Klein & P Li (eds.), *The


First version received: 01.02.2012
Second version received: 03.12.2012
Accepted: 04.12.2012