Sequence of Tense: Syntax, Semantics, Pragmatics

Yael Sharvit

Abstract A theory of embedded tense that derives SOT (Sequence of Tense) effects from an SOT rule is compared with a theory that derives SOT effects without appealing to an SOT rule, and an argument is provided in favor of the former. The argument relies mostly on examples where a tense is embedded under future-in-the-past. Such an argument was originally presented in (Abusch D, Linguist Philos 20:1–50, 1997) and later dismissed in (Altshuler D, Schwarzschild R, Moments of change, cessation implicatures and simultaneous readings. In: Chemla E, Homer V, Winterstein G (eds) Proceedings of Sinn und Bedeutung 17, ENS-Paris, 45–62, 2013a). An additional argument is provided in favor of supplementing the SOT rule with a 'de re' mechanism (as also argued for in Abusch D, Linguist Philos 20:1–50, 1997).

1 The Theoretical Status of the SOT Rule

The English sentence *John said that Mary was self-employed*, where both the matrix and embedded verbs appear in the past tense, can report that John said "Mary is self-employed" (Jespersen 1931). The very same sentence can also report that John said: "Mary was self-employed" (as in *John said that Mary was self-employed in her twenties*). The first reading of the sentence is its null past (or null, for short) reading; the second is its back-shifted reading. The null reading of *John said that Mary was self-employed* resembles the 'de se' reading of *John thinks that he is self-employed*: the latter may report that John thinks "I am self-employed".

Two theories, or families of theories, have been proposed to explain these facts. According to one family of theories, (1) – where ~~past~~ is interpreted as the present from John's perspective – is one of the LFs of *John said that Mary*

Y. Sharvit (✉)
Department of Linguistics, University of California Los Angeles, Los Angeles, CA, USA
e-mail: ysharvit@gmail.com

was self-employed, and (2) – where ~~he~~ is interpreted as first person from John's perspective – is one of the LFs of *John thinks that he is self-employed*.

(1)   John past say [that [. . . ~~past~~ . . .]]   ~~past~~ = "now" (from John's past
                                                          perspective)

(2)   John pres think [that [~~he~~ . . .]]   ~~he~~ = "I" (from John's current
                                                          perspective)

According to the other family of theories, neither (1) nor (2) is generated by the grammar. The null "reading" of *John said that Mary was self-employed* is a special case of its back-shifted reading, and the 'de se' "reading" of *John thinks that he is self-employed* is a special case of its non-'de se' reading.

The LF in (1) arises via the optional application of the SOT rule, which makes past tense invisible to semantic interpretation when embedded under another past tense. My goal is to show that no theory of tense (known to me) can dispense with (1). Since the arguments I present here are applicable only to the tense domain, I take no position with respect to (2). It is quite possible that with existing theories of person pronouns we can – and maybe even should – dispense with (2).

The claim that (1) cannot be dispensed with is based on examples such as (3) – henceforth, the breakfast example – which has a null reading ((3) is constructed after similar examples from Kamp and Rohrer 1983; Abusch 1997).

(3)   A week ago, John <u>decided</u> that in ten days, at breakfast, he <u>would</u> say to his
       mother that they <u>were meeting</u> for the last time.

Crucially, (3) has past tense in the matrix clause, future-in-the-past in its first embedded clause, and past in its most deeply embedded clause. Although it may have other readings, its most salient reading is the one according to which John's plan is to say: "Mom, we are meeting for the last time." Abusch (1997) claims that this reading cannot be accounted for without an SOT rule.

We will see that while the English breakfast example in (3) ***can*** be accounted for without an SOT rule, the existence of languages where the breakfast example has two acceptable counterparts with the same salient reading – one with present in the most deeply embedded clause and one with past in the most deeply embedded clause – is unexpected if such languages have no SOT rule. As we will also see, the following are also problematic for non-SOT theories: (a) variations on (3) with an embedded *before*-clause, and (b) unambiguously back-shifted variations on *John said that Mary was self-employed*.

Regarding the mechanism underlying the SOT rule itself, the literature offers various options (Transmission as in Abusch 1997 and others or as in Grønn and von Stechow 2010 and others, or Deletion as in Ogihara 1996 and others). This paper is agnostic about this issue; we are only concerned with whether the SOT rule – whatever its precise details are – can be dispensed with. For concreteness, we will assume the SOT rule is essentially tense deletion under agreement.

Section 2 presents a failed semantic attempt to dispense with (1). Section 3 presents a pragmatic attempt. Section 4 discusses the shortcomings of the pragmatic

attempt, concluding that (1) cannot be dispensed with. Section 5 discusses some consequences regarding the Upper Limit Constraint (due to Abusch 1997) for any theory of SOT effects.

2 The 'De Re' Theory and Its Failure to Dispense with the SOT Rule

Implicit in the discussion of (1) and (2) above is the assumption that the grammar treats past tense on a par with *he*. Indeed, since Partee 1973, the similarities between tenses and person pronouns have led many tense scholars to treat tenses as non-quantificational (see also Heim 1994; von Stechow 2003; Abusch 1997; Kratzer 1998; Schlenker 1999 and others). However, whether or not tenses are quantificational is orthogonal to the controversy surrounding SOT.[1] So unless otherwise specified (and in the spirit of Prior 1967; Ogihara 1996), tenses are treated here as existential quantifiers over times (which we take to be closed time intervals). Still, we include in our LFs covert pronouns that denote times or worlds and serve as arguments of tenses and verbs (for example, t_0 is a time-denoting pronoun and w_0 a world-denoting pronoun).

Additional basic assumptions include the following. A root node is interpreted relative to an utterance context $c = <W_c, T_c, g_c>$, where W_c is the utterance world, T_c is the utterance time, and g_c is a variable assignment (we ignore additional coordinates such as speaker and hearer merely for simplicity). An utterance context c is suitable for a root LF L only if $[[L]]^c$ is defined, the range of g_c consists of semantic objects that are salient in W_c at T_c, and g_c maps every free occurrence of a time-denoting pronoun in L to T_c and every free occurrence of a world-denoting pronoun in L to W_c. Tenses, which are base-generated as verbal arguments and may move to higher positions, come with covert pronominal arguments that determine an evaluation time (and covert pronominal domain restrictors along the lines of von Fintel 1994, but we omit the latter from the LFs whenever they do not play an essential role). Accordingly, *John said that Mary was self-employed* has at least LF^{PAST} in (4a) as a possible LF, where the free t_0 and the bound t_4 determine the evaluation times of the matrix and embedded clauses, respectively.[2] If English has

---

[1]In addition, on the assumption that pronominal tenses can be bound by a default existential operator (cf. Ogihara 1989), the difference between the two approaches is not so significant anyway.

[2]We also assume the rules Functional Application and Predicate Modification (as in Heim and Kratzer 1998), and:

(a) if α is a pronoun or a trace and i an index: $[[\alpha_i]]^c$ is defined only if $i \in Dom(g_c)$; when defined, $[[\alpha_i]]^c = g_c(i)$.

(b) $[[i\ \alpha]]^c = [\lambda z: [[\alpha]]^{c[i \to z]}$ is defined. $[[\alpha]]^{c[i \to z]}]$, where $c_{[i \to z]}$ is exactly like c except for the possible difference that $g_{c[i \to z]}(i) = z$.

An index i is a pair $<n, \sigma>$, where n is a number and σ a semantic type. We often omit σ to keep the LFs simple.

an SOT rule, the sentence also has LF^PAST in (4b), where the embedded tense is deleted under morphological agreement with the c-commanding main tense (and ~~PAST~~ has no meaning).

(4)   a.   LF^PAST: PAST-t_0 [3 [John say-t_3-w_0 [5 4 [PAST-t_4 [2 [Mary be-t_2-w_5 self-employed]]]]]]
           <u>Back-shifted reading</u>. If c is suitable for LF^PAST, [[LF^PAST]]^c = True iff there is a time t preceding T_c such that for all $<w, t'> \in$ ACC(John, W_c, t), there is a time t'' preceding t' such that Mary is self-employed in w at t''.
      b.   LF^~~PAST~~: PAST-t_0 [3 [John say-t_3-w_0 [5 4 [Mary be-~~PAST~~-t_4-w_5 self-employed]]]]
           <u>Null reading</u>. If c is suitable for LF^~~PAST~~, [[LF^~~PAST~~]]^c = True iff there is a time t preceding T_c such that for all $<w, t'> \in$ ACC(John, W_c, t), Mary is self-employed in w at t'.

In LF^PAST, both tenses move locally, leaving behind traces that are interpreted as bound variables. In LF^~~PAST~~, the matrix tense moves locally leaving behind a trace that is interpreted as a bound variable; the embedded tense is invisible to semantic interpretation, and its time-denoting pronominal argument is interpreted as a bound variable. The meanings of these LFs are derived from the meanings of *PAST* and *say* in (5) and (6) respectively (t'' in (5) is a domain restrictor denoted by a pronoun which, for simplicity, is omitted from (4a,b)).

(5)   For any time t, function p from times to truth values, and time t'',
      [[PAST]]$^c(t'')(t)(p)$ = True iff there is a t' preceding t such that $t' \subseteq t''$ and $p(t')$ = True.
(6)   For any time t, world w, individual x, and function q from worlds to functions from times to truth values:
      [[say]]$^c(t)(w)(q)(x)$ is defined iff ACC(x, w, t) $\neq \emptyset$ and for all $<w', t'> \in$ ACC(x, w, t), $q(w')(t')$ is defined;
      if defined, [[say]]$^c(t)(w)(q)(x)$ = True iff for all $<w', t'> \in$ ACC(x, w, t), $q(w')(t')$ = True.

For current purposes, we make the simplifying assumption that ACC(x, w, t) in (6) – the set of world-time pairs accessible from w and t relative to x – is the set of world-time pairs which, for all x knows/believes in w at t, are the world and time he lives in.

A well-known attempt to derive both readings of *John said that Mary was self-employed* without appealing to an SOT rule is with a 'de re' LF, as in Abusch 1994, 1997 (cf. Enç 1987). The idea is that a 'de re' LF yields roughly the following truth conditions: Regarding some time t that precedes T_c, at some time t' that precedes T_c John said, of t, that Mary's self-employment is within t. This is indeed ambiguous, as t could, in principle, precede t' or be co-temporal with it.

(7)   a.   Null reading T_c

```
----------------------|-----|--------------|----|---------
                              t'
                              t
```

      b.   Back-shifted T_c
           reading

```
----|-----|-------------|-----|--------------|----|---------
     t                  t'
```

Abusch (1997) herself ultimately concluded that the 'de re' attempt is unsuccessful, because of the breakfast example mentioned in Sect. 1, with past-under-future-under-past. For example, the 'de re' LF of *A week ago, John decided that in ten days, at breakfast, he would say to his mother that they were having their last meal together* yields roughly the following truth conditions: With respect to some time t that precedes T_c, at some time t' that precedes T_c by 7 days, John decided, regarding t, that at a time t'' that follows his "now" by 10 days, he says to his mother at t'': We are having – within t – our last meal together. The problem is, roughly, that no time can simultaneously precede T_c and be perceived by John as "now" at a time that he knows is in the future. This suggests that an SOT rule is needed, at the very least to account for the breakfast example.[3] Let us present this reasoning in some more formal detail.

Let us assume – temporarily – that LF^{PAST} in (4b) is not an option made available by the grammar of English. LF^{PAST} in (4a) yields the back-shifted reading only. Assuming that *PAST* can move freely, we obtain – in addition to LF^{PAST} in (4a) – $LF^{PAST\text{-}DERE}$ in (8), where both *PAST*s scope above *say*.

(8) $LF^{PAST\text{-}DERE}$: PAST-t_0 [2 [PAST-t_0 [3 [John say-t_3-w_0 [5 4 [Mary be-t_2-w_5 self-employed]]]]]]

But $LF^{PAST\text{-}DERE}$ in (8) does not yield any attested reading because '4' does not bind any temporal variables.[4] Fortunately, there are more fine-grained versions of 'de re' ascription in the literature (Kaplan 1968; Lewis 1979; Cresswell and von Stechow 1982; Heim 1994; Abusch 1997; Percus and Sauerland 2003) that avoid this problem. Building on Percus and Sauerland 2003; Charlow and Sharvit 2014, we assume the definition of 'time-concept generator' in (9) (where a time-concept is a function from times to functions from worlds to times), and enrich the LF in (8) with the pronoun G_4, which is bound by *say* and whose first argument is t_2 and whose second and third arguments are t_1 and w_5 respectively. $[[[G_4\ t_2]\ t_1]\ w_5]$ in (10) occupies the position of t_2 in (8).

---

[3]See also Heim (1994), Ogihara (1996), von Stechow (1995), Kratzer (1998), Schlenker (1999), Ogihara and Sharvit (2012) and others.

[4]This incorrectly predicts that *John said that Mary was self-employed* cannot be used when John mistakenly located himself at a time that is after T_c and said: "Mary is self-employed now, and was never self-employed before now".

(9)   A time-concept generator suitable for x in world w and time t is a function f such that:

    (a)   the domain of f is the set of times that x is acquainted with in w at t; and

    (b)   for any t' in the domain of f:

        (i) $f(t')$ is a suitable time-concept;

        (ii)   for any $<w', t''> \in$ ACC(x, w, t): $f(t')(t'')(w')$ is defined;

        (iii) $f(t')(t)(w) = t'$.

(10)   Revised LF$^{\text{PAST-DeRE}}$: PAST-t_0 [2 [PAST-t_0 [3 [John say$^{\%}$-t_3-w_0
    [4 5 1 [Mary be-[[[$G_4 t_2$] t_1] w_5]-w_5 self-employed]]]]]]
    If c is suitable for LF$^{\text{PAST-DeRE}}$, [[LF$^{\text{PAST-DeRE}}$]]^c = True iff there are times t
    and t' preceding T_c and a time-concept generator G suitable for John in W_c
    at t' such that for all $<w, t''> \in$ ACC(John, W_c, t'), Mary is self-employed in
    w at $G(t)(t'')(w)$.[5]

A suitable time-concept corresponds to an acquaintance-based description that an attitude holder can use to describe a time to himself. An example of a suitable time-concept is 'now': the smallest function f such that for every t and w, $f(t)(w)$ = t (it is acquaintance-based because we are all acquainted with our "now"). Other examples of suitable time-concepts are 'yesterday' (the smallest function that maps every $<w, t>$ to the day preceding the day surrounding t) and 'the closest time to now that is before now at which the earth moved'. Unsuitable time-concepts are along the lines of 'Sally's birthday' or 'May 3rd, 1952'.

By (9), (i) and (ii) hold for any t and t' preceding T_c and any G suitable for John in W_c at t': (i) if $G(t)$ = 'now', then $G(t)(t')(W_c) = t' = t$, and for all $<w, t''> \in$ ACC(John, W_c, t'), $G(t)(t'')(w) = t'$; (ii) if $G(t)$ = 'yesterday', then $G(t)(t')(W_c)$ (= t) precedes t', and for all $<w, t''> \in$ ACC(John, W_c, t'), $G(t)(t'')(w)$ precedes t''. The option in (i) corresponds to the null reading of *John said that Mary was self-employed* and the option in (ii) corresponds to its back-shifted reading.

This is still not satisfactory, because we have to guarantee that the time-concept generator does not pick out something like 'tomorrow'; otherwise, we predict – counter-intuitively – that *John said that Mary was self-employed* can report that John said: "Mary will be self-employed". According to Abusch, this forward-shifted reading is blocked by the Upper Limit Constraint (ULC), which here we incorporate into a revised definition of 'time-concept generator'.[6]

---

[5]The meaning of *say* in (6) is adjusted as follows: For any p such that p is a function from time-concept generators to functions from world-time pairs to truth values, individual x, time t and world w: [[say$^{\%}$]]$^c(t)(w)(p)(x)$ is defined iff ACC(x, w, t) $\neq$ ø and there is a time-concept generator G suitable for x in w at t such that for all $<w', t'> \in$ ACC(x, w, t), p(G)(t')(w') is defined; if defined, [[say$^{\%}$]]$^c(t)(w)(p)(x)$ = True iff there is a G $\in$ {G*| G* is a time-concept generator suitable for x in w at t and for all $<w', t'> \in$ ACC(x, w, t), p(G*)(t')(w') is defined} such that p(G)(t')(w') = True.

[6]Abusch (1997) contemplates the possibility that the ULC follows from independent principles. In Sect. 5 we mention some attempts (Ogihara 1989, 1996; Bar-Lev 2015; Klecha 2016) to achieve precisely that.

(11)   A time-concept generator suitable for x in world w and time t is a function f such that:
   (a)   the domain of f is the set of times that x is acquainted with in w at t; and
   (b)   for any t' in the domain of f:
       (i) $f(t')$ is a suitable time-concept;
       (ii)   for any $<w', t''> \in \text{ACC}(x, w, t)$: $f(t')(t'')(w')$ is defined;
       (iii) $f(t')(t)(w) = t'$;
       (iv)   for any $<w', t''>$ in the domain of $f(t')$: $f(t')(t'')(w')$ is
               not after t''.   (ULC)

Thanks to the ULC, the time-concept generator can pick out 'now', 'today' and 'yesterday', but not 'tomorrow' or 'next week'. Consequently: (i) the back-shifted reading of *John said that Mary was self-employed* has two sources, the non-'de re' LF$^{\text{PAST}}$, and the 'de re' LF$^{\text{PAST-DERE}}$; and (ii) LF$^{\text{PAST-DERE}}$ also generates a null reading, but no LF generates a forward-shifted reading.

   The theory has two additional important consequences, one positive and the other negative. The positive consequence is its predictions regarding *John says/is saying (now) that Mary was self employed*, with present tense in the matrix clause and past in the embedded clause.

(12)   PAST-t_0 [2 [PRES-t_0 [3 [John say$^\%$-t_3-w_0 [4 5 1 [Mary be-[[[G_4 t_2] t_1] w_5]-w_5 self-employed]]]]]]]
       If c is suitable for (12), $[[(12)]]^c$ = True iff there is a t preceding T_c, a t' overlapping T_c and a time-concept generator G suitable for John in W_c at t' such that for all $<w, t''> \in \text{ACC}(\text{John}, W_c, t')$, Mary is self-employed in w at $G(t)(t'')(w)$.

(13) $[[\text{PRES}]]^c(t'')(t)(p)$ = True iff there is a t' overlapping t such that $t' \subseteq t''$ and $p(t')$ = True.

For any t preceding T_c and t' overlapping T_c and any G suitable for John in W_c at t', (i) and (ii) hold: (i) if $G(t)$ = 'yesterday', it yields a time preceding t'' when applied to any $<w, t''> \in \text{ACC}(\text{John}, W_c, t')$; (ii) if $G(t)$ is something like 'now', it fails to yield t when applied to $<W_c, t'>$. Indeed, the sentence is unambiguously back-shifted (usually, but not always; we come back to this in Sect 4.1).

   The negative consequence is the following. As argued in Abusch 1997, and illustrated by the breakfast example in (3) (with past-under-future-in-the-past), repeated in (14), a 'de re' analysis does not suffice to capture all attested null readings.

(14)   A week ago, John decided that in ten days, at breakfast, he would say to his mother that they were meeting for the last time.

   The null reading of (14) – its most salient reading – is not accounted for. This is the reading according to which John's decision is to say to his mother, at a future time: "Mom, we are meeting (now) for the last time". This led Abusch (along with many other tense scholars) to the conclusion that an SOT rule is needed after all.

As we will see in Sect. 3, Abusch's conclusion has been rejected by proponents of more pragmatic approaches to embedded tenses.

To understand Abusch's conclusion, let us look at some possible LFs of (14). For reasons that do not concern us here, when embedded under a past tense, *would* – which (we assume) is composed of past tense and the future operator *woll* – tends to be interpreted as if its past tense is "null" (implying that John's decision is of the form "I will say to my mother …"). To keep the LFs simple, then, let us ignore the tense morphology of *would*. Accordingly, we derive (15) (where the past tense on *were meeting* moves to a position below *say*), (16) (where the past tense on *were meeting* moves to a position above *say* but below *decide*), and (17) (where the past tense on *were meeting* moves to a position above *decide*). A tense trace is "wrapped" by as many time-concept generators as there are attitude verbs in whose scope the trace is free. Accordingly, the trace of the moved *PAST* in (16) is "wrapped" by one time-concept generator (introduced by *say*), and the trace of the moved *PAST* in (17) is "wrapped" by two time-concept generators (the inner one introduced by *decide*, the outer one introduced by *say*).

(15)   PAST-t_0 [3 [John decide-t_3-w_0 a week ago [5 9 [woll-t_9 in ten days [4 [John say-t_4-w_5 to his mother [7 8 [PAST-t_8 [6 [they be-t_6-w_7 meeting for the last time]]]]]]]]]]]
        If c is suitable for (15), $[[(15)]]^c$ = True iff there is a t that precedes T_c by 7 days such that for all $<w, t'> \in$ ACC(John, W_c, t):
            there is a t'' that is 10 days after t' such that for all $<w', t'''> \in$ ACC(John, w, t''):
                there is a $t^\bullet$ that precedes t''' and John and his mother meet in w' at $t^\bullet$.

(16)   PAST-t_0 [3 [John decide-t_3-w_0 a week ago [5 9 [PAST-t_9 [6 [woll-t_9 in ten days [4 [John say$^\%$-t_4-w_5 to his mother [1 7 8 [they be-[[[G_1 t_6] t_8] w_7]-w_7 meeting for the last time]]]]]]]]]]]]
        If c is suitable for (16), $[[(16)]]^c$ = True iff there is a t that precedes T_c by 7 days such that for all $<w, t'> \in$ ACC(John, W_c, t):
            there is a $t^\bullet$ that precedes t', a t'' that is 10 days after t' and a time-concept generator H suitable for John in w at t'' such that for all $<w', t'''> \in$ ACC(John, w, t''):
                John and his mother meet in w' at $H(t^\bullet)(t''')(w')$.

(17)   PAST-t_0 [6 [PAST-t_0 [3 [John decide$^\%$-t_3-w_0 a week ago [2 5 9 [woll-t_9 in ten days [4 [John say$^\%$-t_4-w_5 to his mother [1 7 8 [they be-[[[G_1 [[[G_2 t_6] t_9] w_5]] t_8] w_7]-w_7 meeting for the last time]]]]]]]]]]]]]
        If c is suitable for (17), $[[(17)]]^c$ = True iff there is a $t^\bullet$ that precedes T_c, a t that precedes T_c by 7 days, and a time-concept generator G suitable for John in W_c at t such that for all $<w, t'> \in$ ACC(John, W_c, t):
            there is a t'' that is 10 days after t' and a time-concept generator H suitable for John in w at t'' such that for all $<w', t'''> \in$ ACC(John, w, t''):
                John and his mother meet in w' at $H(G(t^\bullet)(t')(w))(t''')(w')$.

The non-'de re' LF in (15) is compatible only with a back-shifted reading of the most deeply embedded clause in (14), with John planning to say "Mom, we were meeting for the last time". As shown in (18), neither one of the 'de re' LFs in (16)–(17) allows the most deeply embedded clause in (14) to be null, with John planning to say "Mom, we are meeting for the last time". Likewise, neither one of the 'de re' LFs in (16)–(17) allows the most deeply embedded clause in (14) to be future-oriented, with John planning to say "Mom, we will meet/will be meeting for the last time" (even if the ULC is ignored).

(18) a. Regarding (16): For any t that is 7 days prior to T_c, all $<w, t'> \in$ $\overline{\text{ACC}(\text{John}, \text{W}_c, t)}$, any $t^\bullet$ preceding t', any t'' that is 10 days after t', and any H suitable for John in w at t'':
        if $H(t^\bullet) =$ 'now' or $H(t^\bullet) =$ 'in n days', then $H(t^\bullet)(t'')(w)$ is t'' or some time that is after t''; this is impossible given that t'' is 10 days after t', $t^\bullet$ precedes t' and $H(t^\bullet)(t'')(w) = t^\bullet$.
      b. Regarding (17): For any t that is 7 days prior to T_c, any $t^\bullet$ that precedes $\overline{T_c, \text{all} <w, t'>} \in \text{ACC}(\text{John}, \text{W}_c, t)$, any t'' that is 10 days after t', any H suitable for John in w at t'' and any G suitable for John in W_c at t:
        (i) if $G(t^\bullet) =$ 'in n days', where n $\geq$ 7:
            $G(t^\bullet)(t)(W_c)$ is $t+7$ ($= T_c$) or some later time; this is impossible given that $G(t^\bullet)(t)(W_c) = t^\bullet$ and $t^\bullet$ precedes T_c;
        (ii) if $G(t^\bullet)$ is 'now', 'n days ago', or 'in n days' where $0 < n < 10$:
            if $H(G(t^\bullet)(t')(w)) =$ 'now' or $H(G(t^\bullet)(t')(w)) =$ 'in n days', then $H(G(t^\bullet)(t')(w))(t'')(w)$ is t'' or some time that is after t''; this is impossible given that t'' is 10 days after t', $G(t^\bullet)(t')(w)$ is earlier than $t'+10$, and $H(G(t^\bullet)(t')(w))(t'')(w) = G(t^\bullet)(t')(w)$.

Only two out of these three consequences are welcome. A back-shifted reading of the most deeply embedded clause is pragmatically odd in this particular case, but available in principle, as predicted (and evidenced by the fact that *John decided that in ten days, at breakfast, he would say to his mother that he once loved her* can mean that John is planning to say "Mom, I once loved you"). A forward-shifted reading of the most deeply embedded clause is indeed unavailable, as predicted. Crucially, the fact that no LF yields a null reading of the most deeply embedded clause shows that the grammar must generate the SOT LF in (19), where the embedded *PAST* is deleted under agreement with the matrix *PAST* (and if *would* is represented as *woll + PAST*, that *PAST* is also deleted under agreement with the matrix *PAST*).

(19) PAST-t_0 [3 [John decide-t_3-w_0 [5 9 [woll-(~~PAST~~-)t_9 [4 [John say-t_4-w_5 to his mother [7 8 [they be-~~PAST~~-t_8-w_7 meeting for the last time]]]]]]]]

And of course, once we concede that the breakfast example has an SOT LF, *John said that Mary was self-employed* can also have such an LF, leaving us with two sources for its null reading: the 'de re' LF in (10) and LF$^{\text{PAST}}$ in (4b). *John says that Mary was self-employed* is still correctly predicted not to have a null reading, because the SOT rule – by assumption – applies only under agreement.

Since the non-'de re' LF of *John said that Mary was self-employed* already yields the back-shifted reading and the SOT LF yields the null reading, do we still need the 'de re' mechanism? We do, at the very least in order to account for present-under-past sentences such as *John said that Mary is self-employed*. This sentence does not have a null reading, but only a "double access" reading (see, for example, Smith 1978; Enç 1987), which implies that John said something like "Mary is self-employed this week" (and so, may be unacceptable if the event of saying occurred in the distant past; for example, given that pregnancies don't last long, *Twenty years ago, John said that Mary is pregnant* is infelicitous). Assuming that the English *PRES* – unlike *PAST* – can be bound only as a consequence of the application of the SOT rule (von Stechow 1995; Ogihara 1996), *John says that Mary is self-employed* and *John said that Mary is self-employed* receive the SOT LF in (20) and the 'de re' LF in (21), respectively, but neither of them can receive the non-'de re' LF in (22).[7]

(20) PRES-t_0 [3 [John say-t_3-w_0 [5 2 [Mary be-~~PRES~~-t_2-w_5 self-employed]]]]

(21) PRES-t_0 [2 [PAST-t_0 [3 [John say$^\%$-t_3-w_0 [4 5 9 [Mary be-[[[G_4 t_2] t_9] w_5]-w_5 self-employed]]]]]]
      If c is suitable for (21), $[[(21)]]^c$ = True iff there is a t overlapping T_c, a t' preceding T_c and a time-concept generator G suitable for John in W_c at t' such that for all $<w, t''> \in$ ACC(John, W_c, t'), Mary is self-employed in w at $G(t)(t'')(w)$.

(22) *PAST/PRES-t_0 [3 [John say-t_3-w_0 [5 2 [PRES-t_2 [3 [Mary be-t_3-w_5 self-employed]]]]]]

By (11), any time-concept generator introduced in (21) picks out a time-concept such as 'this week' or 'today'. This means that $G(t)(t'')(w)$ overlaps t'' and $G(t)(t')(W_c)$ overlaps t' (for any relevant G, t, t', t'' and w). Furthermore, since $G(t)(t')(W_c) = t$, t overlaps both T_c and t', and self-employment holds throughout $G(t)(t'')(w)$ (being a stative predicate, *be self-employed* has the sub-interval property; see, for example, Dowty 1977).

Thus, according to this view, the complete theory of SOT effects – which we refer to from now on as SOT+'de re' – has an SOT rule, a 'de re' mechanism, and a rule that rules out (22).

A final remark. Not all embedded predicates induce SOT effects. The generalization seems to be (see Stowell 2007) that only stative-like predicates (i.e., purely stative adjectives such as *self-employed*, as well as complex predicates containing certain operators such as the progressive) support null readings. This requirement – which we call the stativity requirement – is illustrated by the contrast in (23): the non-stative verb *leave* (when appearing without the progressive operator or the habitual operator) is non-stative-like.

---

[7][PAST/PRES-t_0 [3 [John say-t_3-w_0 [5 6 [PAST/PRES-t_0 [3 [Mary be-t_3-w_5 self-employed]]]]]]], where '6' does not bind anything, is ruled out by whatever principle rules out (8).

(23) a.   John said that Mary was self-employed.   Stative
           May report that John said: "Mary is self-employed."
      b.   John said that Mary left.   Non-stative
           May <u>not</u> report that John said: "Mary leaves/is leaving".
      c.   John said that Mary was leaving.
           May report that John said: "Mary is leaving"   +PROG

The stativity requirement does not follow from anything we have said so far. However, in Sect. 4 we will see that the stativity requirement is sometimes lifted anyway, so it is not so clear that we even want to make it part of the theory of SOT effects anyway.

The next section discusses another attempt to dispense with the SOT rule. This attempt builds the stativity requirement into the theory of SOT effects.

3   The Scalar Theory

The scalar theory due to Altshuler and Schwarzschild (2013a, b) dispenses not just with LF^PAST in (4b), as an LF of *John said that Mary was self-employed*, but also with LF^PAST-DE RE in (10). Accordingly, *John said that Mary was self-employed* has only one semantic reading, which arises from the back-shifted (non-'de re') LF of the sentence – its only LF – and which may undergo pragmatic strengthening. The option of pragmatic strengthening makes the sentence "ambiguous". Pragmatic strengthening often results in a cessation implicature (which is a scalar implicature). Lack of pragmatic strengthening may lead to the perception – or illusion – of a null reading.

The scalar theory aims to explain the correlation between the stativity requirement (illustrated in (23) above) and the explicit mention of cessation/non-cessation in unembedded sentences. This correlation is illustrated in (24). (We note, but do not discuss any further, the fact that it is less obvious that activity predicates, which are not stative-like, show the expected pattern; cf. the unambiguously back-shifted *John said that Mary pushed a cart*).

(24) a.   Mary was self employed; she no longer is. / She still is.   cf. (23a)
      b.   Mary left; *she no longer does/is. / *She still does/is.   cf. (23b)
      c.   Mary was leaving; she no longer is / She still is.   cf. (23c)

The acceptability of both variants of (24a) suggests that the inference that self-employment has ceased (see Klein 1994; Musan 1997 and others) is a cancellable implicature of *Mary was self-employed* rather than an entailment. By contrast, the unacceptability of all variants of (24b) suggests that the inference that Mary's leaving is over is an entailment of *Mary left*, rather than an implicature. The correlation between (23) and (24) further suggests that an illusion of a null reading of embedded clauses arises only when there is no implicature that forces cessation by the local evaluation time. The challenge for such a view is to predict *John says/is saying that Mary was self-employed*, with present in the matrix clause and past in the embedded clause, to be incompatible with a situation where John says "Mary is

self-employed". Here is an attempt (which may not be entirely faithful to Altshuler and Schwarzschild 2013a, b)[8].

According to Gricean approaches to scalar implicatures, upon hearing a sentence, we tend – in the absence of conflicting information – to infer that all stronger scalar alternatives to that sentence (if there are any) are false, because – we reason – the speaker has made the strongest claim she can commit to. There are various ways to implement this idea; here we use grammatical exhaustification along the lines of Fox 2007 (this choice is driven only by the semantic transparency that grammatical exhaustification affords). Accordingly, *Mary is reading a book or a newspaper* is interpreted as 'Mary is reading a book or a newspaper but not both' whenever *exh* is present at LF because, by assumption, *and* is a lexical alternative to *or*.

(25) For any suitable utterance context c, [0 7 [Mary read-t_7-w_0 a book and a newspaper]] $\in$ ALT([0 7 [Mary read-t_7-w_0 a book or a newspaper]], c), therefore, if defined:
[[exh [0 7 [Mary read-t_7-w_0 a book or a newspaper]]]]^c(W_c)(T_c) = True only if [[0 7 [Mary read-t_7-w_0 a book and a newspaper]]]^c(W_c)(T_c) = False.

(26) For any α and utterance context c such that [[α]]^c is a function from worlds to functions from times to truth values:

a. for any world w and time t: [[exh α]]^c(w)(t) is defined iff c is suitable for α and [[α]]^c(w)(t) is defined. When defined, [[exh α]]^c(w)(t) = True iff [[α]]^c(w)(t) = True and for all $\beta \in$ ALT(α, c), [[β]]^c(w)(t) = False.

b. ALT(α, c) = {β| β is a syntactic alternative to α and a negatable alternative to α relative to c}.

c. for any β:
(i) β is a syntactic alternative to α iff β is derivable from α by replacing at least one node γ in α with one of γ's lexical alternatives.
(ii) β is a negatable alternative to α relative to c iff {<w, t>| [[β]]^c(w)(t) is defined} $\neq \emptyset$ and {<w, t>| [[β]]^c(w)(t) = True} $\subset$ {<w, t>| [[α]]^c(w)(t) = True}.

By assumption, *exh* is inserted unless its insertion results in a contradiction, as is the case when the potential implicatures are explicitly "cancelled" (for example, *Mary is reading the book or the newspaper*; *in fact, she is reading both*).

To account for cessation implicatures along these lines, we make certain additional assumptions. We assume, as before, that c is suitable for root L only if [[L]]^c is defined, and g_c maps any free occurrence of a time-denoting pronoun in L to T_c and any free occurrence of a world-denoting pronoun in L to W_c. We also assume that the time-concept 'now' (which maps every <w, t> to t) is salient in any suitable c (i.e., there is at least one index n such that g_c(n) = 'now').

---

[8]Altshuler (2016) discusses some further implications of the scalar theory that are not addressed here. It is also worth noting that the proposal in Gennari 2003, like the scalar theory, does not acknowledge an SOT LF, but unlike the scalar theory it recognizes null and back-shifted readings as two semantically distinct readings.

In addition, we assume that the English present tense has the meaning in (27) (modified from Altshuler and Schwarzschild 2013b), which implies that it is partly "absolute" (or indexical) and partly "relative"; f is a suitable time-concept that serves here as a tense domain restrictor. The past tense has the more standard "relative" meaning in (28).

(27) $[[\text{PRES}^{\text{AS}}]]^c(f)(t)(w)(p)$ is defined iff: (i) there is a t' at or after T_c such that $t' \subseteq f(T_c)(W_c)$, (ii) $t \subseteq f(t)(w)$, and (iii) for all $t' \subseteq f(t)(w)$: $p(t')$ is defined. If defined, $[[\text{PRES}^{\text{AS}}]]^c(f)(w)(t)(p) = $ True iff for all $t' \subseteq f(t)(w)$: $p(t') = $ True.

(28) $[[\text{PAST}^{\text{AS}}]]^c(f)(t)(w)(p)$ is defined iff there is a $t' \subseteq f(t)(w)$ such that t' precedes t and $p(t')$ is defined.
       If defined, $[[\text{PAST}^{\text{AS}}]]^c(f)(t)(w)(p) = $ True iff there is a $t' \in \{t'' | t'' \subseteq f(t)(w), t''$ precedes t and $p(t'')$ is defined$\}$ such that $p(t') = $ True.

The entries in (27)–(28) mimic Abusch's predictions regarding "double access" sentences such as *John said that Mary is self-employed* (see Sect. 2), without a 'de re' LF, when f1 and f8 have the same value; e.g., 'this week' (by default, the embedded restrictor is anaphoric to the matrix restrictor).

(29)   (i) PAST^{AS}-f_1-t_0-w_0 [3 [John say-t_3-w_0 [5 6 [PRES^{AS}-f_8-t_6-w_5 [4 [Mary be-t_4-w_5 self-employed]]]]]]

       (ii) c is suitable for (i) only if: there is a t at or after T_c such that $t \subseteq g_c(8)(T_c)(W_c)$, and there is a $t \subseteq g_c(1)(T_c)(W_c)$ such that t precedes T_c and for all $<w, t'> \in \text{ACC}(\text{John}, W_c, t)$, $t' \subseteq g_c(8)(t')(w)$.
       If c is suitable for (i), $[[(\text{i})]]^c = $ True iff there is a $t^* \in \{t| t \subseteq g_c(1)(T_c)(W_c)$, t precedes T_c, and for all $<w, t'> \in \text{ACC}(\text{John}, W_c, t)$, $t' \subseteq g_c(8)(t')(w)\}$ such that for all $<w, t> \in \text{ACC}(\text{John}, W_c, t^*)$, for all $t' \subseteq g_c(8)(t)(w)$, Mary is self-employed in w at t'.

In addition, we assume that stative-like predicates have, in addition to the sub-interval property, the temporal profile of statives in (30). Accordingly, stative events may have a lower bound and/or an upper bound, but have no starting point or endpoint.

(30)   The temporal profile of statives:[9]
       For any tenseless stative clause S, if S is true at moment m, then there is a moment m1 preceding m and a moment m2 following m such that S is true at [m1,m2] (the time that begins at m1 and ends at m2).

---

[9]This is modified from Altshuler and Schwarzschild 2013a. See Altshuler 2016 for discussion of the progressive and the temporal profile of statives.

Finally, we assume that for any index n and temporal restrictor f, $\{\text{PAST}^{AS}\text{-}f_m|\ m$ is an index$\}$ is the set of lexical alternatives to $\text{PRES}^{AS}\text{-}f_n$, and $\{\text{PRES}^{AS}\text{-}f_m|\ m$ is an index$\}$ is the set of lexical alternatives to $\text{PAST}^{AS}\text{-}f_n$.

The emergence of a cessation implicature is illustrated in (31) for *Mary was self-employed*.

(31) a. [exh [0 7 $[\text{PAST}^{AS}\text{-}f_1\text{-}t_7\text{-}w_0$ [3 [Mary be-t_3-w_0 self-employed]]]]]-w_0-t_7

   b. Let $\alpha = $ [3 [Mary be-t_3-w_0 self-employed]].
      For any utterance context c suitable for (31a), and any n such that
      $g_c(n) = $ 'now' and $\{<w, t>|\ [[\text{PRES}^{AS}\text{-}f_n\text{-}t_7\text{-}w_0\ \alpha]]^{c[0\rightarrow w,7\rightarrow t]} = \text{True}\}$
      $\subset \{<w, t>|\ [[\text{PAST}^{AS}\text{-}f_1\text{-}t_7\text{-}w_0\ \alpha]]^{c[0\rightarrow w,7\rightarrow t]} = \text{True}\}$:

      (i) by (26b) and (27), [0 7 $[\text{PRES}^{AS}\text{-}f_n\text{-}t_7\text{-}w_0\ \alpha]] \in \text{ALT}([0\ 7$
          $[\text{PAST}^{AS}\text{-}f_1\text{-}t_7\text{-}w_0\ \alpha]], $ c), therefore,

      (ii) by (26a), $[[(31a)]]^c = \text{True}$ only if $[[\text{PRES}^{AS}\text{-}f_n\text{-}t_7\text{-}w_0\ \alpha]]^c = $
          False.

```
[                  ] = self-employment            T_c
-----|--|----------------------------------------------|-------|--->
      past_self-employ
[_____] =
g_c(1)(T_c)(W_c)
```

It follows that there is at least one non-self-employment time in T_c, whenever $[[(31a)]]^c = \text{True}$ and there are negatable alternatives.

Since *exh* is inserted by default, and since the semantics of *be self-employed* does not entail cessation, there are two types of circumstances where *Mary was self-employed* does not implicate that Mary is no longer self-employed, as specified in (32).

(32) (i) The cessation implicature is explicitly "cancelled", so *exh* is not inserted.

   (ii) *exh* is inserted but there are no relevant negatable alternatives.

Explicit "cancellation" is illustrated by *Mary was self-employed; in fact, she still is*. Vacuous exhaustification is illustrated by *I entered the room; there was a book on the table; the book was in Russian*, which (usually) does not imply that the book is no longer in Russian (cf. Klein 1994). The reason is, presumably, that c is set up in a way that gurantees that for all n, $[[f_n]]^c$ is past-oriented (i.e., for all w and t in Dom($[[f_n]]^c$), $[[f_n]]^c(t)(w)$ precedes t) or $[[f_n]]^c = $ 'now', which means that for all m, $\{<w, t>|\ [[\text{PRES}^{AS}\text{-}f_m\text{-}t_7\text{-}w_0$ [3 [the book be-in-Russian-t_3-w_0]] $]]^{c[0\rightarrow w,7\rightarrow t]}$ is defined$\} = \emptyset$, or $\{<w, t>|\ [[\text{PRES}^{AS}\text{-}f_m\text{-}t_7\text{-}w_0$ [3 [the book be-in-Russian-t_3-w_0]] $]]^{c[0\rightarrow w,7\rightarrow t]} = \text{True}\} \not\subset \{<w, t>|\ [[\text{PAST}^{AS}\text{-}f_1\text{-}t_7\text{-}w_0$ [3 [the book be-in-Russian-t_3-w_0]] $]]^{c[0\rightarrow w,7\rightarrow t]} = \text{True}\}$.

In addition, if $PAST^{AS}$ is replaced with $PRES^{AS}$, ALT([0 7$[\text{PRES}^{AS}\text{-}f_1\text{-}t_7\text{-}w_0\ \alpha]], $ c) $= \emptyset$, because $PRES^{AS}$ has no negatable alternatives. Furthermore, by the
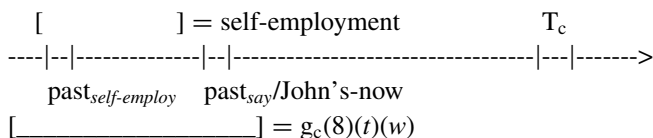
semantics of $PRES^{AS}$, *Mary is self employed* entails that Mary is self-employed at T_c. If *be self-employed* is replaced with a predicate such as *leave*, which lacks the temporal profile of statives, for any n such that $g_c(n)$ = 'now', $\{<w, t>|$ [[$PRES^{AS}$-f_n-t_7-w_0 [3 [Mary leave-t_3-w_0]]]]$^{c[0\rightarrow w, 7\rightarrow t]}$ = True} $\not\subset$ $\{<w, t>|$ [[$PAST^{AS}$-f_1-t_7-w_0 [3 [Mary leave-t_3-w_0]]]]$^{c[0\rightarrow w, 7\rightarrow t]}$ = True}, so exhaustification never results in negating *Mary leaves/is leaving*. Furthermore, by the semantics of *leave*, *Mary left* entails – rather than implicates – that one event of Mary leaving begins and ceases pre-T_c (though there could be subsequent leaving events, pre-T_c, at T_c and post-T_c).

Since *be self-employed* is stative-like, when *Mary was self-employed* is embedded under *John said/says that* . . . , cessation of self-employment by John's "now" at the saying time is never entailed, but may be implicated. The exhaustified LFs of *John said that Mary was self-employed* and *John says/is saying that Mary was self-employed* are given in (33); (34) illustrates how exhaustification works in each case: in at least one of the accessible worlds, self-employment cannot go on uninterrupted throughout John's "now".

(33)   a.   <u>past-under-past:</u>
            [exh [0 7 [$PAST^{AS}$-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5 6
            [$PAST^{AS}$-f_8-t_6-w_5 [2 [Mary be-t_2-w_5 self-employed]]]]]]]]]-w_0-t_7

        b.   <u>past-under-present:</u>
            [exh [0 7 [$PRES^{AS}$-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6
            [$PAST^{AS}$-f_4-t_6-w_5 [2 [Mary be-t_2-w_5 self-employed]]]]]]]]]-w_0-t_7

(34)   Let α = [2 [Mary be-t_2-w_5 self-employed]].
        a.   For any utterance context c suitable for (33a), and any n such that
            $g_c(n)$ = 'now' and $\{<w, t>|$ [[$PAST^{AS}$-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5
            6 [$PRES^{AS}$-f_n-t_6-w_5 α]]]]]]$^{c[0\rightarrow w, 7\rightarrow t]}$ = True} $\subset$ $\{<w, t>|$
            [[$PAST^{AS}$-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [$PAST^{AS}$-f_8-t_6-w_5
            α]]]]]]$^{c[0\rightarrow w, 7\rightarrow t]}$ = True}:
            (i)   by (26b), (27)-(28) and the semantics of *say*, [0 7
                [$PAST^{AS}$-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [$PRES^{AS}$-f_n-t_6-w_5
                α]]]]]] $\in$ ALT([0 7 [$PAST^{AS}$-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5 6
                [$PAST^{AS}$-f_8-t_6-w_5 α]]]]]], c), therefore
            (ii)   by (26a), [[(33a)]]^c = True only if [[$PAST^{AS}$-f_1-t_7-w_0 [3 [John
                say-t_3-w_0 [5 6 [$PRES^{AS}$-f_n-t_6-w_5 α]]]]]]^c = False.

```
    [                    ] = self-employment              T_c
----|--|--------------|--|----------------------------|---|------->
    past_self-employ   past_say/John's-now
    [_____] = g_c(8)(t)(w)
```

b. For any utterance context c suitable for (33b), and any n such that $g_c(n)$ = 'now' and $\{<w, t>|$ [[PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PRES^{AS}-f_n-t_6-w_5 α]]]]]]$^{c[0\rightarrow w, 7 \rightarrow t]}$ = True$\} \subset \{<w, t>|$ [[PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PAST^{AS}-f_4-t_6-w_5 α]]]]]]$^{c[0\rightarrow w, 7 \rightarrow t]}$ = True$\}$:

  (i) by (26b), (27)-(28) and the semantics of *say*, [0 7 [PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PRES^{AS}-f_n-t_6-w_5 α]]]]]] $\in$ ALT([0 7 [PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PAST^{AS}-f_4-t_6-w_5 α]]]]]], c), therefore

  (ii) by (26a), [[(33b)]]^c = True only if [[PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PRES^{AS}-f_n-t_6-w_5 α]]]]]]^c = False.

self-employment = [   ] T_c
----------------------------|---|------------------------|----|------------------->
                                  past$_{self\text{-}employ}$ pres$_{say}$/John's-now
                            [_____] = $g_c(4)(t)(w)$

An illusion of a null reading may arise when exhaustification is vacuous. The closer the time introduced by the embedded past tense to John's "now" at the saying time is, the easier it is to obtain an illusion of a null reading.

In (33a), an illusion of a null reading easily arises because in many suitable contexts there are no relevant negatable syntactic alternatives. For example, if for all n, [[f_n]]^c is past-oriented or [[f_n]]^c = 'now', then for any m, [0 7 [PAST^{AS}-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PRES^{AS}-f_m-t_6-w_5 α]]]]]] $\notin$ ALT([0 7 [PAST^{AS}-f_1-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PAST^{AS}-f_8-t_6-w_5 α]]]]]], c); cf. (32(ii)). On the other hand, an illusion of a null reading cannot arise in (33b), in practice, for the following reason. For any c, [[(33b)]]^c = True only if $g_c(9)$ is present-oriented and $g_c(4)$ is compatible with past tense. By default, the embedded tense restrictor is anaphoric to the matrix tense restrictor; that is to say, $g_c(9) = g_c(4)$. A possible value for both, then, is along the lines of 'today'. Given this, [0 7 [PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PRES^{AS}-f_n-t_6-w_5 α]]]]]] $\in$ ALT([0 7 [PRES^{AS}-f_9-t_7-w_0 [3 [John say-t_3-w_0 [5 6 [PAST^{AS}-f_4-t_6-w_5 α]]]]]], c) for all n such that $g_c(n)$ = 'now', and self-employment cannot go on uninterrupted throughout John's "now".

To sum up, on the assumption that *exh* is inserted by default: (a) in past-under-past sentences with embedded stative-like predicates, an illusion of a null reading easily arises because *exh*-insertion can easily result in no cessation; and (b) in past-under-present sentences with embedded stative-like predicates, an illusion of a null reading rarely – in practice, perhaps never – arises, because the embedded tense restrictor is anaphoric to the present-oriented matrix tense restrictor (and cessation is guaranteed whenever *exh* is inserted).

Importantly, an illusion of a null reading may arise not only in the simplex past-under-past *John said that Mary was self-employed*, but also in the complex past-

under-past breakfast example – *A week ago, John decided that in ten days he would say to his mother that they were meeting for the last time*. The LF that underlies this breakfast example is (35).

(35) $PAST^{AS}$-f_1-t_{10}-w_0 [3 [John decide-t_3-w_0 [5 9 [$PAST^{AS}$-f_{12}-t_9-w_5 [6 [woll-t_6 [4 [John say-t_4-w_5 to his mother [7 8 [$PAST^{AS}$-f_{11}-t_8-w_7 [2 [they be-t_2-w_7 meeting for the last time]]]]]]]]]]]]

On the assumption that both *woll* and *be meeting* have the temporal profile of statives, an illusion of a null reading easily arises when c is such that (35) has no relevant negatable alternatives (in which case we may infer that John's decision was: "I will say to my mother: we are meeting for the last time").

The scalar theory has an additional positive consequence that is worth pointing out: it does not appeal to the ULC to guarantee that *John said that Mary was self-employed* has no forward-shifted reading. This simply follows from the assumption that the only LFs available for this sentence are not 'de re' LFs, but rather LFs with an embedded quantificational past.

We now compare SOT+'de re' with the scalar theory, ruling in favor of the former.[10]

4   In Defense of SOT+'de re'

We take issue with some of the main claims of the scalar theory, namely: (i) that a potential lack of cessation obviates the need for an SOT rule, (ii) that a partly "relative" English present tense obviates the need for 'de re' LFs, (iii) that the scalar theory predicts a correlation between explicit mention of cessation/non-cessation in main clauses and null readings in embedded clauses, and (iv) that stativity is a pre-requisite for the emergence of null readings in embedded clauses. As we show, SOT+'de re' makes the following predictions: (a) it predicts there to be SOT languages, non-SOT languages and "mixed" languages, challenging (i) and (ii) and vindicating Abusch's claim that no theory of SOT effects is complete without an SOT rule or a 'de re' mechansim (see Sect. 4.1), (b) it predicts the absence of a null reading in certain past-under-past sentences with stative-like predicates, challenging (iii) (see Sect. 4.2), and (c) it predicts null readings of breakfast examples with non-stative-like predicates, challenging (iv) (see Sect. 4.3).

---

[10]A concern regarding the scalar theory arises with respect to vacuous exhaustification of $\neg\alpha$. Cessation is not guaranteed in negated past-under-present sentences (e.g., *John isn't saying that Mary was self-employed*) and in embedded polar questions (e.g., *John is asking whether Mary was self-employed*). However, since negation also poses some non-trivial problems for 'de re' LFs (as discussed in Charlow and Sharvit 2014 and others), we refrain from comparing SOT+'de re' with the scalar theory on the basis of their predictions regarding negation.

4.1   Non-SOT Languages and "Mixed" Languages

Not all languages express a null reading in the same way. SOT+'de re' predicts the attested language typology straightforwardly, but the scalar theory does not.

There are languages where the most natural way to report that someone said "Mary lives in Paris" is with an embedded present. One instance of such a language is Modern Hebrew (see Sharvit 2003; Ogihara and Sharvit 2012, and others). The most natural way to report a past event of Dan saying "Mira lives in Paris" is with the present-under-past (36) (which need not imply that Mira's living is ongoing at T_c, according to Dan), and the most natural way to report a past event of Dan saying "Mira lived in Paris" is with the past-under-past (37).

(36)   Dan   amar   Se   Mira   mitgoreret   be-pariz
       Dan   say-past   COMP   Mira   live-pres   in-Paris
       Dan: "Mira lives in Paris"

(37)   Dan   amar   Se   Mira   hitgorera   be-pariz
       Dan   say-past   COMP   Mira   live-past   in-Paris
       Dan: "Mira lived in Paris"

According to SOT+'de re', this is explained as follows. English has an SOT rule and its present tense cannot be bound unless the SOT rule has applied (see (20)–(22)). Hebrew lacks an SOT rule, and its present tense has the same semantics as *PRES* (see (13)), but unlike the English *PRES*, the Hebrew $PRES^H$ can be bound when embedded. This explains why (36) has a null reading (with no obligatory "double access" effect).

(38)   a. $[[PRES^H]]^c = [[PRES]]^c$
       b.   LF of (36) (cf. (22)): PAST-t_0 [8 [Dan say-t_8-w_0 [5 6 [$PRES^H$-t_6 ...]]]]

This makes three additional predictions. Firstly, in the breakfast sentence in (39), a null reading of the most deeply embedded clause, according to which the subject describes an event ongoing in his "now", is available only with an embedded present tense; an embedded past tense supports only a back-shifted reading (see Sharvit 2003, 2008; Hatav 2012; this is also true of breakfast examples in Japanese, see Ogihara 1996; Ogihara and Sharvit 2012).[11]

(39)   lifney   Savua, Dan xaSav   Se   Miri hayta   amura   tox asara
       before   week   Dan think-past COMP Miri be-past supposed in ten
       yamim   lomar le-ima   Se   hi   mitgoreret/*hitgorera be-pariz
       days   to-say to-mother-hers COMP she live-pres/live-past   in-Paris

       Mira: "Mom, I live in Paris".

---

[11]I thank Moysh Bar-lev and Idan Landau for sharing with me their native speaker judgments regarding (36), (37) and (39).

Secondly, (37) does sometimes have a null reading (though subject to speaker variation). This is because a 'de re' LF of (37) is available, which is similar to the 'de re' LF of *John said that Mary was self-employed* in (10) in Sect. 2 (see Sharvit 2003, 2008, Ogihara and Sharvit 2012; Bar-Lev 2015). Importantly, a 'de re' LF cannot rescue the past variant of (39), for the same reason that it does not account for its English counterpart, as we saw in Sect. 2.

Thirdly, there are languages where a null reading arises under past, either with past or with present in the embedded clause (see Sharvit 2003). This is because SOT+'de re' does not ban languages that have both the SOT rule and a locally bindable present. Indeed, in Modern Greek (unlike English and Hebrew), both past and present can be null under past. This is illustrated by the Greek counterpart of the breakfast examples in (40), where a null reading is available with both past and present.[12],[13]

(40) Prin   mia evdomada o   Jorghos ipe   oti   se
      before one week   the   Jorghos sayPFV.past-3SG.   that in
      dheka meres tha eleghe   stin   kopela   tu   oti
      ten   days FUT past.IMPF.say-3SG. to-the girlfriend of-his that
      sinadjodusan/sinadjunte   ja   teleftea fora
      meet.pastIMPV-3PL/meet.pres-3PL   for   last   time

      Jorghos: "We are meeting for the last time"

Within the scalar theory the Hebrew present tense has no "absolute" component; it is purely "relative" as in (41a); this explains why (36) has a null reading (with no obligatory "double access" effect).

(41) a. $[[\text{PRES}^h]]^c(f)(t)(w)(p)$ is defined only if $t \subseteq f(t)(w)$ and $p(t)$ is defined.
          If defined, $[[\text{PRES}^{h2}]]^c(f)(t)(w)(p)$ = True iff there is a $t' \in \{t'' | t'' \subseteq$
          $f(t)(w)$ and $p(t'')$ is defined$\}$ such that $p(t')$ = True.
      b. LF of (36) (cf. (29)): PAST^{AS}-f_1-t_0-w_0 [3 [Dan say-t_3-w_0 [5 6
          [PRES^h-f_8-t_6-w_5 ...]]]]

Accordingly, English constitutes a clear violation of The Embeddability Principle (Sharvit 2003), which says (or, more accurately, implies) that every language should be able to express the proposition expressed by the Hebrew (36). The reason is this: since the English grammar lacks a purely "relative" present (and, like all languages, also lacks an SOT rule), it has the means to deliver a proposition that is compatible

---

[12]I thank Nikos Angelopoulos, Cleo Condoravdi, Anastasia Giannakidou, Sabine Iatridou, Pinelopi Koujianou Goldberg, Eleni Miltsakaki and Orest Xherija for sharing with me their native speaker judgments regarding (40).

[13]The modal *tha* in (40) does not have a null reading when accompanied by present tense, but this is probably due to the way it interacts with aspect (see Iatridou 2000). Regardless, the embedded verb corresponding to *meet* is good with either past or present (on the "null" reading).

with the proposition expressed by (36), but does not have the means to deliver that proposition itself (because semantically, it delivers only "double access" readings of the English counterpart of (36) and only back-shifted readings of the English past-under-past counterpart of (36)). We now question this consequence.

To account for the fact that the present-variant of both (39) and (40) supports a null reading, we have to assume that both Hebrew and Greek have $PRES^h$. In order to make the right prediction regarding (39) (namely, that only its present-variant supports a null reading), we must somehow ensure that the set of relevant negatable syntactic alternatives to the past-variant of (39) (i.e., alternatives whose negation causes cessation of living in Paris) is never empty. On the other hand, in order to make the right prediction regarding the Greek (40) (namely, that both present- and past-variants support a null reading), we have to ensure that the set of relevant negatable alternatives to the past-variant of (40) (i.e., alternatives whose negation causes cessation of meeting) may be empty. While it is not impossible to imagine a theory that "blames" cross-linguistic variation on whether the set of relevant alternatives may be empty, such a solution would be incompatible with (37). This suggests that the breakfast example does, after all, show that some languages have an SOT rule, that the null reading is a semantic reading in its own right, and that the Embeddability Principle is respected in English (though this may require a minor adjustment of (13)). Granted, SOT+'de re' must recognize – and explain – the emergence of a cessation inference in, for example, *Mary was self-employed*, but this does not conflict with recognizing that past-under-past sentences may have a genuine semantic null reading.

In addition, the fact that the Hebrew (37) sometimes has a null reading illustrates the need for 'de re' LFs independently of English present-under-past sentences (which, as we saw in Sect. 3, receive an alternative explanation within the scalar theory). The scalar theory cannot account for the fact that (37) may have a null reading: any mechanism for generating syntactic alternatives that guarantees cessation of living in Paris for the past-variant of (39) guarantees cessation of living in Paris for (37).[14]

It is worth noting that English, too, provides some evidence for 'de re' LFs. Consider the past-under-present in (42). The received wisdom is that it is unambiguously back-shifted.

(42)   John says/is saying that Mary was self-employed.

My own impression, based on the responses of the speakers I consulted, is that it indeed never has a pure null reading, but some speakers allow – what I call – a pseudo-null reading in very special circumstances. By "pseudo-null reading" I mean: a reading that implies that John is misinformed about his temporal location

---

[14]The Greek counterpart of (37) has a null reading too, as expected on the assumption that Greek has an SOT rule. The Japanese counterpart of (37) does not have a null reading at all; this is unexpected on the assumption that it lacks an SOT rule but has a 'de re' LF. See Ogihara and Sharvit (2012) for a possible explanation.

(as people sometimes are; see Lewis 1979; Cresswell and von Stechow 1982), and mistakenly takes the past to be his "now" by relying on an unreliable source of information to determine his temporal location. Suppose John wakes up from a coma in a hospital, and he has just met Mary, who is on the phone talking to clients. He does not know what the date is, so he looks at the calendar on the wall. The calendar is showing January, despite the fact that it is February (because the busy hospital staff neglected to update it). John, for whom the calendar is the only source of information regarding his temporal location, says: "Mary is self-employed." While, admittedly, many speakers do not accept *John is saying that Mary was self-employed* as a faithful report of this, some speakers do.

Within SOT+'de re', an SOT LF is ruled out – due to lack of agreement between the matrix and embedded tenses – but the 'de re' LF (43) is generated ((12) in Sect. 2).

(43) PAST-t_0 [2 [PRES-t_0 [3 [John say$^\%$-t_3-w_0 [4 5 1 [Mary be-[[[G_4 t_2] t_1] w_5]-w_5 self-employed]]]]]]]

If c is suitable for (43), $[[(43)]]^c$ = True iff there is a t preceding T_c, a t' overlapping T_c and a time-concept generator G suitable for John in W_c at t' such that for all $<w, t''> \in$ ACC(John, W_c, t'), Mary is self-employed in w at $G(t)(t'')(w)$.

Time-concepts such as 'now' will not work for (43), as we saw in Sect. 2. But there are other options to consider. When John is misinformed because the calendar shows January instead of February, it seems that speakers vary with respect to the range of time-concepts they admit. Some speakers admit a time-concept that assigns to every t and w the time that is presented to John as "now" in w at t. This means that $G(t)(t'')(w)$ overlaps t'', and $G(t)(t')(W_c)$ is t (for all relevant G, t, t', t'' and w), which precedes T_c. For other speakers, such a time-concept fails to meet some rigorous definition of the term "suitable".

The scalar theory predicts a cessation implicature for (42), with (33b) in Sect. 3 as its default LF, because the embedded tense restrictor is anaphoric to the matrix tense restrictor. With respect to those speakers who do not permit pseudo-null readings of (42), this is a welcome prediction. However, if we want to generate such readings (and account for those speakers who do permit these readings), we must stipulate that in a context where John relies on a "bad" source to determine his temporal location, *exh* is not inserted (or the anaphoricity requirement is suspended). The problem is that we would predict no contrast between a past-induced mistake and a future-induced mistake, contrary to fact. Imagine we are watching a video of John and reporting what is happening, in real time, to someone who has no visual or auditory access to the video. While we can use (42) to report that John is saying "Mary is self-employed" after determining his temporal location by looking at a calendar that presents last week as this week, we cannot use it to report that John is saying "Mary is self-employed" after determining his temporal location by looking at a calendar that presents next week as this week. This is not predicted, if exhaustification can be avoided whenever John uses a "bad" source. By contrast,

the 'de re' LF in (43) makes the right prediction: $G(t)(t')(W_c)$ is t itself – the time shown by the calendar – and it cannot, by the ULC, follow t'.

4.2 Unambiguous Sentences

Sentences where the embedded verb is overtly modified by a past-oriented modifier are unambiguously back-shifted. This suggests that (non-)cessation and the (un-)availability of null readings are independent of each other.

Recall that the scalar theory aims to account for the correlation between the stativity requirement of null readings in embedded clauses (illustrated by the fact that a null reading is available for *John said that Mary was sad* but unavailable for *John said that Mary left*; cf. (23)) and explicit mention of cessation/non-cessation in main clauses (illustrated by the contrast between *Mary was sad and she still is* and **Mary left and she still does/is*; cf. (24)). That correlation breaks down when the relevant verb phrase has a past-oriented overt temporal modifier (see (44)–(45)) and/or past perfect morphology (see (46)–(47)).

(44)   Mary was sad two minutes ago. She still is. / She no longer is.

(45)   a.   A minute ago, John said that Mary was sad two minutes before.
                 John: "Mary was sad two minutes ago".
         b.   A minute ago, John said that Mary was sad (then).
                 John: "Mary is sad (now)".
                 John: "Mary was sad (then)."

(46)   Mary finally went to bed. She had been upset, but she no longer was/and she still was.

(47)   a.   John said that Mary had been upset half an hour before.
                 John: "Mary was upset half an hour ago."
         b.   John said that Mary had always been upset.
                 John: "Mary has always been upset."
         c.   John said that Mary was upset (then).
                 John: "Mary is upset (now)".
                 John: "Mary was upset (then)".

When faced with the task of choosing between (45a) and (45b) to report that John uttered a minute ago, "Mary is sad", speakers invariably choose the latter. And when asked, "What did John say?" after being presented with (45a) and (45b), speakers never respond with "Mary is sad" to the (45a)-option (but may respond in this manner to the other option). Similarly, when faced with the task of choosing between (47a), (47b) and (47c) to report that John uttered, "Mary is upset", speakers invariably choose (47c). And when asked, "What did John say?" after being presented with (47a), (47b) and (47c), speakers never respond with "Mary is upset" to the (47a,b)-options (but may respond in this manner to the other option).

The fact that explicit mention of cessation/non-cessation is possible in (44) and (46), despite the temporal modifier/past perfect morphology, suggests that non-cessation and null readings are distinct phenomena.

Indeed, within SOT+'de re', an SOT LF is available for (45b), but ruled out for (45a), because a deleted tense (which picks out John's "now") conflicts with the past-oriented *two minutes before* (which picks out a time prior to John's "now"). A 'de re' LF is ruled out for (45a) whenever the time-concept picked out by the concept generator conflicts with the past-oriented *two minutes before* (e.g., when the concept-generator picks out the time-concept 'now'). This correctly predicts (45a) to be back-shifted. Similarly, an SOT LF is available for (47b) (but not for (47a)) with the implication that John uttered, "Mary has always been upset" (which is not the same as "Mary is upset"), and for (47c).

Notice that the fact that cessation/non-cessation can be explicitly mentioned in (48) is compatible with these predictions: back-shifted readings are compatible both with cessation and with non-cessation (on either theory).

(48)   A minute ago, John said that Mary was/had been sad two minutes before, and that she still was / and that she no longer was.

The predictions of the scalar theory depend on what we take the syntactic alternatives to be. It seems plausible to assume that those alternatives may include lexical alternatives to the past tense, to the overt temporal modifiers and to the aspect markers. There are certainly many contexts suitable for (45a) in which negation of some of these syntactic alternatives results in cessation of sadness (by John's "now"). But by the same reasoning that predicts (33a) to sometimes lack relevant negatable alternatives, as well as (45b), there are also many suitable contexts in which (45a) lacks relevant negatable alternatives, and we expect the illusion of a null reading to be available. Similarly, there are many suitable contexts in which (47a) lacks relevant negatable alternatives. Any mechanism that would enforce relevant negatable alternatives in (45a) or (47a) would be ad-hoc and/or at odds with the fact that cessation is not enforced in either (44) or (46).

Given this, it perhaps comes as no surprise that stativity is not always a pre-requisite for null readings, as we show in Sect. 4.3.

4.3   SOT Effects and Non-Stative-Like Predicates

As we saw, stativity is often a pre-requisite for a null reading of a past-under-past sentence. However, the stativity requirement is lifted in certain breakfast examples.

Consider the examples in (49) and (50). In (49a), with past tense on *graduate*, graduation occurs after moving but pre-T_c (Anscombe 1964; Ogihara 1996 and others). In (49b), there is an apparent mismatch between *sold* and *tomorrow* (Dudman 1983, 1984; Iatridou 2000; Ippolito 2003 and others). In (50a), with present tense on *graduate*, graduation still occurs after moving, but it is not co-temporal with T_c. In (50b), selling is not co-temporal with T_c either.

(49)   a.   John moved before he graduated.
       b.   If John sold his stocks tomorrow, he would make a lot of money.
(50)   a.   John will move before he graduates.
       b.   If John sells his stocks tomorrow, he will make a lot of money.

Interestingly, when we embed (49a) under *John said that he would* … and
(49b) under *John said today that* …, as in the (51), they "acquire" a future-
in-the-present reading that involves John uttering "… I will …" (see (53)),
making a neutral prediction about his future, in addition to the future-in-the-past
reading that involves John uttering "… I would …" (see (52)), possibly making
counterfactual/subjunctive claims.

(51)   a.   John said that he would move before he graduated.
       b.   John said today that if he sold his stocks tomorrow, he would make a lot
            of money.
(52)   a.   Future-in-the-past reading of (51a)
            John said: "I would move before I graduated".
       b.   Future-in-the-past reading of (51b)
            John said today: "If I sold my stocks tomorrow, I would make a lot of
            money."
(53)   a.   Future-in-the-present reading of (51a)
            John said: "I will move before I graduate".
       b.   Future-in-the-present reading of (51b)
            John said today: "If I sell my stocks tomorrow, I will make a lot of
            money."

The paraphrases of the future-in-the-present readings of (51a) and (51b) show that
they are null readings with respect to *woll*, *graduate* and *sell*. Given that *woll* has the
temporal profile of statives, it is conceivable that an illusion of a null reading with
respect to *woll* arises because of this profile. But *graduate* and *sell* both lack the
temporal profile of statives. How do the null readings arise with respect to *graduate*
and *sell*, then?

    Let us assume that the meaning of *before* is along the lines of Krifka 2010 (cf.
Anscombe 1964), and that tenses are doubly-indexed time-denoting pronouns that
may be bound by *before*, *woll*, ∃, etc.[15]

---

[15]We use Krifka's theory of *before* because it is advocated in Altshuler and Schwarzschild 2013a,
but our point is the same within other theories of *before* (e.g., Condoravdi 2010). The decision
to switch to pronominal tenses is dictated by Krifka's *before*, which introduces an existential
quantifier that binds tenses. In addition, we treat the present as fully "relative" for simplicity, but
our point remains the same if we make it partly "absolute" as in Altshuler and Schwarzschild 2013b
(cf. (27)).

(54) a. $[[\text{before}]]^c(p)(t) = \text{True}$ iff there is no t' at or prior to t such that $p(t') =$
         True.

      b. For any c such that $g_c(i)$ and $g_c(j)$ are times:
         $[[\text{past}_{i,j}]]^c = g_c(i)$ (where $g_c(i)$ precedes $g_c(j)$);
         $[[\text{pres}_{i,j}]]^c = g_c(i)$ (where $g_c(i)$ overlaps $g_c(j)$).

The LFs in (55)–(56) account for the intuitive meanings of (49a) and (50a,b) without
an SOT rule. Following Kratzer (1986), we assume that an *if*-clause restricts the
quantifier that appears in the consequent clause of a conditional; in (50b), that
quantifier is future-in-the-present, namely, the future operator *woll* with present
tense. We also assume that a *before*-clause and its sister combine by Predicate
Modification.

(55) (49a): $\exists$ [[1 [John move-past$_{1,0}$-w$_0$]] [before 3 [he graduate-past$_{3,0}$-w$_0$]]]
      (Moving precedes graduation.)

(56) a. (50a): $\exists$ 2 [woll-pres$_{2,0}$ [[1 [John move-pres$_{1,1}$-w$_0$]] [before 3 [he
         graduate-pres$_{3,3}$-w$_0$]]]]
         (Moving and graduation are post-T$_c$, and moving precedes graduation.)

      b. (50b): $\exists$ 2 [[woll-pres$_{2,0}$ [1 [John sell-pres$_{1,1}$-w$_0$ tomorrow]]] [1 [he
         make-t$_1$-w$_0$]]]
         (Selling and making money are post-T$_c$.)

The sentence in (49b) has a subjunctive LF – not provided here – which resolves
the apparent mismatch between *sold* and *tomorrow*. The subjunctive future-in-the-
past readings of (51a,b) come from subjunctive LFs as well, whatever their precise
details might be.

    The problem is with the future-in-the-present readings of (51a,b), whose LFs are
not subjunctive LFs (recall that according to those readings, John makes neutral
predictions about his future). Without an SOT rule, we may obtain (57a) for (51a)
and (57b) for (51b).

(57) a. $\exists$ 3 [John say-past$_{3,0}$-w$_0$ [9 5 [$\exists$ 2 [woll-past$_{2,5}$ [[1 [he move-t$_1$-w$_9$]]
         [before 1 [he graduate-past$_{1,5}$-w$_9$]]]]]]]

      b. $\exists$ 3 [John say-past$_{3,0}$-w$_0$ today [9 5 [$\exists$ 2 [[woll-past$_{2,5}$ [1 [he
         sell-past$_{1,5}$-w$_9$ tomorrow]]] [1 [he make-t$_1$-w$_9$]]]]]]

Thanks to the assumed stativity of *woll*, we may obtain an illusion of a null reading
with respect to *woll* (cf. (35)), predicting the inference that John uttered "... I
will ...". However, the readings described in (53) are still not accounted for. (57a),
which implies that graduation may follow John's "now", is compatible with gradua-
tion preceding moving, because the *before*-clause fails to restrict post-John's-"now"
graduations. But intuitively, moving and graduating may occur post-John's-"now".
It is reasonable to assume that (57a) is ruled out on pragmatic grounds (as its truth
conditions are very weak). (57b) implies that selling both precedes and follows
John's "now", at least whenever John is well-informed about his temporal location,

and is therefore ruled out (crucially, the future-in-the-present reading of (51b) does not depend on John being mistaken about his temporal location).

In (57a,b), the "evaluation time" index of the embedded past tense is bound by *say*. What if it were bound by the embedded '∃', as in (58)?

(58)   a.   ∃ 3 [John say-past$_{3,0}$-w$_0$ [9 5 [∃ 2 [woll-past$_{2,5}$ [[1 [he move-t$_1$-w$_9$]]
             [before 1 [he graduate-past$_{1,2}$-w$_9$]]]]]]]]
       b.   ∃ 3 [John say-past$_{3,0}$-w$_0$ today [9 5 [∃ 2 [[woll-past$_{2,5}$ [1 [he
             sell-past$_{1,2}$-w$_9$ tomorrow]]] [1 [he make-t$_1$-w$_9$]]]]]]]

The LF in (58a) is completely uninformative; it says that no pre-moving graduation can be post-moving. (58b) may yield an attested and informative reading of (51b), but it must be banned because it incorrectly predicts *If John sold his stocks tomorrow, he will make a lot of money* to be well-formed. If the "evaluation time" index of the embedded past is not bound at all, informative readings depend on John being mistaken about his temporal location. But neither future-in-the-present reading in (51a,b) depends on such a mistake.

On the other hand, if the grammar of English has an SOT rule, which deletes one tense index under agreement (and, by assumption, $[[\text{past}_i]]^c = [[\text{pres}_i]]^c = g_c(i)$), the future-in-the-present readings are predicted straightforwardly: (59a), an SOT-variant of (57a), yields the reading paraphrased in (53a); (59b), an SOT-variant of (57b), yields the reading paraphrased in (53b).

(59)   a.   ∃ 3 [past$_{3,0}$ [4 [John say-t$_4$-w$_0$ [9 5 [∃ 2 [woll-past$_2$ [[1 [he move-t$_1$-w$_9$]]
             [before 1 [he graduate-past$_1$-w$_9$]]]]]]]]]]
       b.   ∃ 3 [past$_{3,0}$ [4 [John say-t$_4$-w$_0$ today [9 5 [∃ 2 [[woll-past$_2$ [1 [he
             sell-past$_1$-w$_9$ tomorrow]]] [1 [he make-t$_1$-w$_9$]]]]]]]]]

Importantly, once we acknowledge the SOT rule, SOT LFs of (49a) and (50a,b) also become available, in principle (cf. Heim 1994; Ogihara 1996; Sharvit 2014).

(60)   (49a): ∃ 4 [past$_{4,0}$ [[1 [John move-t$_1$-w$_0$]] [before 3 [he
       graduate-past$_3$-w$_0$]]]]
(61)   a.   (50a): ∃ 2 [pres$_{2,0}$ [4 [woll-t$_4$ [[1 [John move-t$_1$-w$_0$]] [before 3 [he
             graduate-pres$_3$-w$_0$]]]]]]
       b.   (50b): ∃ 2 [pres$_{2,0}$ [4 [[woll-t$_4$ [1 [John sell-pres$_1$-w$_0$ tomorrow]]]
             [1 [he make-t$_1$-w$_0$]]]]]]

This allows us to ban identical indices on tenses (cf. (22)), thereby ruling out the LFs in (56) and the LF in (62), correctly predicting the unacceptability of *John moved before he graduates* (in English, the corresponding Japanese sentence is acceptable; see Ogihara 1996; Sharvit 2014).

(62)   *∃ [[1 [John move-past$_{1,0}$-w$_0$]] [before 3 [he graduate-pres$_{3,3}$-w$_0$]]]

Notice that while the SOT LFs in (61) yield the intuitive truth conditions of (50a,b), the SOT LF in (60) yields truth conditions for (49a) that are too weak (see Kubota et al. 2011). This suggests the following (see Sharvit 2014): (a) that *say* and *woll* are SOT triggers, but *before* is not; (b) that (59a,b) (where SOT is triggered by *say/woll*) are the LFs of (51a,b), and (61a,b) are the LFs of (50a,b) (where, again, SOT is triggered by *say/woll*).[16]

Notice also that even in the subjunctive LF of (51a) (whatever its precise details might be), the past tense of *graduate* must be "deleted", as shown by the fact that (63a) (with a "non-deleted" past) allows graduation to precede moving in the counterfactual worlds quantified over by the modal (cf. (57a)), but (63b) (with a "deleted" past) requires moving to precede graduation in those counterfactual worlds.

(63) a. ... [9 5 [MODAL ... [6 [[1 [he move-t_1-w_6]] [before 1 [he graduate-past$_{1,5}$-w_6]]]]]]
   b. ... [9 5 [MODAL ... [6 [[1 [he move-t_1-w_6]] [before 1 [he graduate-past$_1$-w_6]]]]]]

To sum up, while SOT+'de re' does not readily explain why the stativity requirement is lifted when *before* or *if* intervenes between past tense and *say* (or, for that matter, why the stativity requirement is there in the absence of *before/if*), it does generate the future-in-the-present readings of (51a,b), while a theory that lacks an SOT rule does not. The fact that null tenses do not always obey the stativity requirement undermines any theory that "blames" null readings on the temporal profile of statives, rather than on an SOT rule which, on its own, does not discriminate between predicates that are stative-like and predicates that are not.

5   Summary and some Remarks on the ULC

We compared the SOT+'de re' theory of embedded tense with the scalar theory and saw that while basic SOT effects are explained by both theories, complex embeddings seem to favor SOT+'de re', especially with regard to cross-linguistic variation, unambiguously back-shifted sentences, and the stativity requirement. Taken together, these facts vindicate Abusch's claim that no theory of embedded tense is complete without an SOT rule and a 'de re' mechanism.

---

[16]The waitress at the Los Angeles bar where parts of this paper were written once asked me: "Did you want anything from the happy hour menu before it ended?". While it is not entirely clear why *do* is in-the-past, the fact that *end* is in-the-past is expected, on the assumption that *want*, like *say*, triggers SOT.

As we saw, the ULC plays a central role in the 'de re' component of SOT+'de re'. Some tense scholars have argued that the effects of the ULC follow from more general principles and some have argued that the ULC is empirically inadequate. My own position, motivated below, is that the ULC is empirically adequate, though it is certainly desirable to derive its effects from general principles. The criticism of the scalar theory in Sect. 4 is independent of whether the ULC is an independent principle or a derived principle.

An interesting attempt to derive ULC effects from general principles is made in Klecha 2016, where the ban on forward-shifted readings of past-under-past sentences is taken to be verb-dependent rather than absolute. According to Klecha, some embedding verbs are associated with a modal base that is incompatible with forward-shifted readings, and some are associated with a modal base that is compatible with forward-shifted readings. As Klecha reports, in certain dialects of English, *John hoped that Mary was self-employed* has a forward-shifted reading. Crucially, the predictions of the ULC-based theory – regarding *say*-type verbs – match Klecha's predictions. Therefore, even if Klecha is right, and ULC effects follow from general principles, it is harmless – in the context of this paper – to talk about the ULC as if it were a principle in its own right.

A completely different approach is adopted in Bar-Lev 2015, where it is claimed that the ULC over-generates, even with *say*-type verbs. Bar-Lev observes that a past-under-past sentence with a universal quantifier in subject position does not have a mixed reading. For example, *Every man said that Mary was unemployed* has a null reading, according to which every man said "Mary is unemployed", and a back-shifted reading, according to which every man said: "Mary was unemployed". What the sentence does not have, according to Bar-Lev, is a reading according to which some man/men said "Mary is unemployed" and some man/men said "Mary was unemployed." As shown in (64), such a reading is predicted by the ULC-based 'de re' theory, because the time-concepts can vary with men. Since G_4 is bound in the scope of *every man*, some men could, in principle, be associated with the time-concept 'now' and some men with 'last month'. As shown in (65), a mixed reading is also predicted by the scalar theory, by "skolemizing" the embedded tense restrictor. Since $PAST^{AS}$-$[f_4$ $pro_6]$ is in the scope of *every man* and pro_6 is bound by *every man*, some men may be associated with cessation of unemployment, and some men with non-cessation of unemployment.

(64)   every man-w_0 [8 [PAST-t_0 [2 [PAST-t_0 [3 [t_8 say$^{\%}$-t_3-w_0 [4 5 7 [Mary be-[[[G_4 t_2] t_7] w_5]-w_5 unemployed]]]]]]]]

(65)   every man-w_0 [6 [[exh [0 9 [PAST^{AS}-$[f_1$ $pro_6]$-t_9-w_0 [3 [t_6 say-t_3-w_0 [5 7 [PAST^{AS}-$[f_4$ $pro_6]$-t_7-w_5 [2 [Mary be-t_2-w_5 unemployed]]]]]]]]]]-w_0-t_9]]

Bar-Lev adopts the view in Ogihara 1989, 1996, where the ULC is not recognized as a principle; rather, ULC effects follow from the way tense traces are interpreted in complement clauses of attitude verbs. Bar-Lev strengthens Ogihara's restrictions on tense traces so that 'de re' past-traces and present-traces never have a shifted reading, backward or forward (this, in Bar-lev's system, follows from the Copy Theory of Movement, along the lines of Fox 2002). Importantly, even if Bar-Lev's

observation regarding mixed readings is correct, this has no effect on the claims made here regarding the necessity to adopt an SOT rule in order to account for the facts discussed in Sect. 4. Even in Bar-Lev's system, accounting for these facts requires the application of an SOT rule.

It seems to me that *Every man said that Mary was unemployed* usually does not have a mixed reading, but I'm not convinced it never does. Consider *Every man said, on a different occasion, that Mary was unemployed during the month on the calendar in front of him* and suppose all the men who said "Mary is unemployed this month" are amnesiacs who rely on the calendars in front of them to figure out their temporal location (and it is possible that the calendars are not up-to-date). In such a state of affairs the sentence is acceptable (to some speakers, at least) even if there is also a recovered amnesiac who said "Mary was unemployed last month".[17]

Whereas Bar-Lev claims that the ULC is too weak, Altshuler and Schwarzschild (2013b) claim that it is too strong because, they say, a "double access" reading of present-under-past sentences is not obligatory (see also Bary and Altshuler 2014). This is based on the following Air Berlin baggage counter exchange, which can only be coherent if *The stewardess told me you have my bags* has a forward-shifted reading.

(66) Customer: I believe you have my bags.
      Employee: Who said I have your bags?
      Customer: The stewardess told me you have my bags.
      Employee: When did she tell you that?
      Customer: On the flight.

Since *The stewardess told me you have my bags* cannot, given the context, have a "double access" reading (during the flight, presumably, both the passenger and his bags were on the plane), its acceptability suggests that the ULC is not part of the semantics of this sentence. Indeed, according to Altshuler and Schwarzschild 2013b, the "double access" effect of present-under-past sentences is context-dependent; it arises when $PRES^{AS}$ is interpreted in the embedded clause, and need not arise when $PRES^{AS}$ is raised to the matrix clause.

I maintain that the ULC is needed, even if we concede that (66) is coherent (though there seems to be some speaker variation). This is based on the strangeness of the (made up) exchanges in (67), with *think* and *ask*, and on the contrasts in (68), which show that a forward-oriented temporal adverbial is incompatible with present-under-past ((68c) shows that, in principle, *now* could cover the telling time, but this option is pragmatically unavailable if it implies that the bags were at the baggage counter while the customer was still flying).

---

[17]Admittedly, this sentence requires speakers to be receptive to such a reading to begin with; as we saw in Sect. 4.1, in connection with past-under-present sentences, many speakers are not.

(67) a. Customer: I believe you have my bags.
          Employee: Why do you think I have your bags?
          Customer: Because the stewardess <u>thought</u> that you have my bags.
          Employee: When did she think that?
          Customer: #On the flight.
      b. Customer: I believe you have my bags.
          Employee: Why do you think I have your bags?
          Customer: I <u>asked</u> the stewardess whether you have my bags and
                     she said Yes.
          Employee: When did you ask her that?
          Customer: #On the flight.
(68) a. Two hours ago, on the flight, the stewardess told me that you would
          have my bags now.
      b. #Two hours ago, on the flight, the stewardess told me that you have
          my bags now.
      c. Ten minutes ago, I ran into the stewardess and she told me that you
          have my bags now.

The strangeness of (67a,b) and the contrasts in (68) indicate that there is something
special about (66). I would like to suggest that *The stewardess told me that you
have my bags* in (66) is a case of pseudo-'de re' (discussed in Kaplan 1977 and
elsewhere). Consider (69), and assume that Bill and Jane are in a public area; Bill is
a graduate student and Jane is not.

(69) Announcement: Graduate students should see the registrar.
      Jane to Bill: They just said that <u>you</u> should see the registrar.

   The announcer did not utter "Bill should see the registrar". In fact, it stands to
reason that the announcer was not even acquainted with Bill during the announce-
ment (and they were not in the same room even when the announcement was made).
The acceptability of *They just said that you should see the registrar* seems to be due
to the reasoning in (70).

(70) a. Given what the announcement said, if we were to say to the announcers:
          "This is Bill; he is a graduate student; should he see the registrar?", they
          would respond: "Yes".
      b. Therefore, they sort of said that Bill should see the registrar.

If this reasoning is supported by the context, *They just said that you should see
the registrar* receives a special 'sort of said'-interpretation, via a special LF, where
you is not "wrapped" by a concept-generator. Notice that in the same context,
They think that you should see the registrar is odd. Likewise, in a context where
someone asks Jane "Should graduate students see the registrar?", she would need
a very good reason to felicitously say to Bill *These people asked me whether you
should see the registrar* (normally, this would be a very odd thing for her to ask).
This suggests that the concept-generator-less LF is not the default LF (though

what exactly the restrictions on such an LF are is not entirely clear). Similarly, a concept-generator-less LF may support a pseudo 'de re' interpretation of present tense, as illustrated in (71).

(71) a. If I had said to the stewardess, at 2pm: "Suppose it's 5pm and I am at the baggage counter; do the baggage people have my bag?", she would have responded: "Yes".
      b. Therefore, the stewardess sort of told me, at 2pm, that the baggage counter people have my bag (now, at 5pm).

It seems, then, that the ULC is empirically valid, at least in the verb-dependent sense of Klecha 2016, if not in the absolute sense of Abusch 1997.

References

Abusch, D. (1994). Sequence of tense revisited: two semantic accounts of tense in intensional contexts. In H. Kamp (Ed.), *Ellipsis, Tense and Questions, Dyana 2 Deliverable*, 87–139.

Abusch, D. (1997). Sequence of tense and temporal de re. *Linguistics and Philosophy, 20*, 1–50. doi:10.1023/A:1005331423820.

Altshuler, D. (2016). *Events, states and times*. Warsaw: De Gruyter.

Altshuler, D., & Schwarzschild, R. (2013a). Moments of change, cessation implicatures and simultaneous readings. In E. Chemla, V. Homer, & G. Winterstein (Eds.), *Proceedings of Sinn und Bedeutung 17* (pp. 45–62). ENS-Paris.

Altshuler, D., & Schwarzschild, R. (2013b). Correlating cessation with double access. In M. Aloni, M. Franke, & F. Roelofsen (Eds.), *Proceedings of the 19th Amsterdam Colloquium* (pp. 43–50).

Anscombe, G. E. (1964). Before and after. *The Philosophical Review, 73,* 3–24. doi:10.2307/2183199.

Bar-Lev, M. E. (2015). De-re tenses and trace conversion. In *Proceedings of SALT 25*.

Bary, C., & Altshuler, D. (2014). Double access. In *Proceedings of Sinn und Bedeutung 19*.

Charlow, S., & Sharvit, Y. (2014). Bound 'de re' pronouns and the LFs of attitude reports. *Semantics and Pragmatics 7*, Article 3, 1–43. http://dx.doi.org/10.3765/sp.9.14

Condoravdi, C. (2010). NPI licensing in temporal clauses. *Natural Language and Linguistic Theory, 28*, 877–910.

Cresswell, M., & von Stechow, A. (1982). 'De re' belief generalized. *Linguistics and Philosophy, 5*, 503–535. doi:10.1007/BF00355585.

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

Dudman, V. H. (1983). Tense and time in English verb clusters of the primary pattern. *Australian Journal of Linguistics*, 3, 25–44

Dudman, V. H. (1984). Conditional interpretation of *if*-sentences. *Australian Journal of Linguistics, 4*, 143–204.

Enç, M. (1987). Anchoring conditions for tense. *Linguistic Inquiry, 18*, 633–657.

Fox, D. (2002). Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry, 33*(1), 63–96. doi:10.1162/002438902317382189.

Fox, D. (2007). Free choice disjunction and the theory of scalar implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics* (pp. 71–120). New York: Palgrave-Macmillan.

Gennari, S. (2003). Tense meaning and temporal interpretation. *Journal of Semantics, 20*, 35–71.

Grønn, A., & von Stechow, A. (2010). Complement tense in contrast: The SOT parameter in Russian and English. In A. Grønn, and I. Marijanovic (Eds.), Russian in Contrast. *Oslo Studies in Language 2*(1), 245–267.

Grønn, A., & von Stechow, A. (2011). Future vs. Present in Russian and English adjunct clauses. *Scando-Slavica, 57*(2), 109–153.

Hatav, G. (2012). Bound tenses. In R. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 638–668). Oxford: Oxford University Press.

Heim, I. (1994). Comments on Abusch's theory of tense. In H. Kamp (Ed.), *Ellipsis, tense and questions*, *DYANA deliverable R2.2.B* (pp. 143–170). Amsterdam: University of Amsterdam.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Journal of Linguistic Inquiry, 31*, 231–270. doi:10.1162/002438900554352.

Ippolito, M. (2003). Presuppositions and implicatures in counterfactuals. *Natural Language Semantics, 11*, 145–186.

Jespersen, O. (1931). *A Modern English Grammar on Historical Principles, Part IV, Syntax, third volume, Time and Tense*. Heidelberg: Carl Winters Universitatsbuchhandlung.

Kamp, H., & Rohrer, C. (1983). Tense in texts. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language* (pp. 250–269). Berlin: de Gruyter.

Kaplan, D. (1968). Quantifying in. *Synthese, 19*(1), 178–214.

Kaplan, D. (1977). Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives. Published in J. Almog, J. Perry, and H. Wettstein (Eds.), *Themes from Kaplan*, 1989 (pp. 481–614). Oxford: Oxford University Press.

Klecha, P. (2016). Modality and embedded temporal operators. *Semantics and Pragmatics, 9*, 1–55.

Klein, W. (1994). *Time in language*. London: Routledge.

Kratzer, A. (1986). Conditionals. *Chicago Linguistics Society, 22*(2), 1–15.

Kratzer, A. (1998). More structural analogies between pronouns and tenses. In *Proceedings of SALT*, VIII. CLC Publications, Cornell University.

Krifka, M. (2010). Before and after without coercion. *Natural Language and Linguistic Theory, 28*, 911–929.

Kubota, Y., Lee, J., Smirnova, A., and Tonhauser, J. (2011). Cross-linguistic variation in temporal adjunct clauses. In *Cahier Chronos: Selected Proceedings of Chronos 8*. Rodopi, Amsterdam/Atlanta.

Lewis, D. (1979). Attitudes de dicto and de se. *The Philosophical Review, 88*, 513–543.

Musan, R. (1997). Tense, predicates, and lifetime effects. *Natural Language Semantics, 5*(3), 271–301.

Ogihara, T. (1989). *Temporal reference in English and Japanese*. University of Texas at Austin PhD dissertation.

Ogihara, T. (1996). *Tense, attitudes and scope*. Dordrecht: Kluwer Academic Publishers.

Ogihara, T., & Sharvit, Y. (2012). Embedded tenses. In R. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 638–668). Oxford: Oxford University Press.

Partee, B. (1973). Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy, 70*(18), 601–609.

Percus, O., & Sauerland, U. (2003). On the LFs of attitude reports. In *Proceedings of Sinn und Bedeutung 7* (pp. 228–242).

Prior, A. (1967). *Past, present and future*. Oxford: Clarendon Press.

Schlenker, P. (1999). *Propositional attitudes and indexicality*. PhD dissertation, MIT.

Sharvit, Y. (2003). Embedded tense and universal grammar. *Linguistic Inquiry, 34*, 669–681.

Sharvit, Y. (2008). The puzzle of free indirect discourse. *Linguistics and Philosophy, 31*, 353–395 doi:10.1007/s10988-008-9039-9.

Sharvit, Y. (2014). On the universal principles of tense embedding: The lesson from *before*. *Journal of Semantics, 31*, 263–313. doi:10.1093/jos/ffs024

Smith, C. (1978). The syntax and interpretation of temporal expressions in English. *Linguistics and Philosophy, 2*, 43–99. doi:10.1007/BF00365130.

von Fintel, K. (1994). *Restrictions on quantifier domains*. PhD dissertation, University of Massachusetts at Amherst.

von Stechow, A. (1995). On the proper treatment of tense. In *Proceedings of SALT 6*.

von Stechow, A. (2003). *Feature deletion under semantic binding: Tense, person, and mood under verbal quantifiers*. Text of the NELS33 talk.

Stowell, T. (2007). The syntactic expression of tense. *Lingua, 117*, 437–463.