NOTES ON TENSE INTERPRETATION[1]

Tue Trinh, MIT 2007

In the Priorian theory of tense (Prior 1967, Montague 1973), time is an evaluation parameter, just like worlds and assignments. Formatives in the object language that affect temporal interpretation of sentences are not themselves meaningful expressions, but are defined syncategorematically.[2] For example, the morpheme *PAST* in (1a) has the definition in (1b).

(1) a. PAST [John dance]
b. $[\![\text{PAST } \phi]\!]^{w,t,g} = 1$ iff there is a time t' before t such that $[\![\phi]\!]^{w,t',g} = 1$

Calculation of the truth condition for (1a) at a world w and a time t will give (2).

(2) There is a time t' before t such that John dances at t' in w

This accords with our intuition. Another prediction of this system is that when a past tense sentence is embedded under another past tense sentence, we get the so called "backward shifted" reading. Consider sentence (3a) with LF (3b).

(3) a. John said Bill kissed Mary
b. PAST [John say [PAST [Bill kiss Mary]]]

Calculation of the truth condition of (3b) with respect to a world w, a time t and an assignment g will yield (4).

(4) There is a time t' before t such that for all w' compatible with what John says at t' in w, there is a time t'' before t' such that Bill kisses Mary at t'' in w'

Again, this seems to be correct. For (3a) to be true, it seems that the alleged kissing of Mary by Bill must have occurred before John's report. If what John said in the past is "Bill will kiss Mary", then (3a) cannot be uttered truthfully.

It has been noted that as a matter of empirical fact, not every "past under past" construction has the backward shifted reading. Consider example (5).

(5) Hillary married the man who became president

Intuitively, (5) is true if either Hillary's husband had become president before the marriage, or Hillary's husband would become president after the marriage. The Priorian theory can give us the first reading but not the second, which is called the "later-than-matrix" reading. Thus, (5) has the LF in (6a), and calculation of the truth condition for (6a) with respect to a world w, a time t and an assigment g will yield (6b). No ambiguity is predicted.[3]

---

[1] I thank Danny Fox and Irene Heim for valuable discussion and helpful comments on this squib.

[2] These two points are independent. We can have a system in which time is an evaluation parameter and tense operators are meaningful expressions.

[3] I follow Enç (1987) and assume that the temporal interpretation of nominal predicates is not semantic, i.e. not index sensitive. In what follows, I will treat nouns as having no tense.

(6) a. PAST [Hillary marry [the man [who$_1$ PAST t$_1$ become president]]]
b. There is a time t' before t such that Hillary marries at t' in w an individual x such that there is a time t'' before t' and x becomes president at t'' in w

One solution to this problem that has been proposed makes use of QR (Ladusaw 1977, Stowell 1993). Specifically, it is assumed that the object NP can (type-shift and) raise at LF above the matrix PAST operator, as in (7).

(7) [the man [who$_1$ PAST t$_1$ become president]]$_2$ [PAST Hillary marry t$_2$]

The PAST operator in the relative clause is outside the scope of the matrix PAST operator in (7). Consequently, the former is not interpreted with respect to the time introduced by the latter, and no backward shifted reading results. Calculation of the truth condition for (7) with respect to a time t, a world w and an assignment g will yield the following.

(8) There is a time t' before t s.t. Hillary marries at t' in w the individual x such that x is a man in w and there is a time t'' before t and x becomes president at t'' in w

Call this the scope analysis. We now turn to another problem for the Priorian theory. Consider (9).

(9) John said Bill liked Mary

For this sentence to be true, what John said at some point in the past can be either "Bill liked Mary" or "Bill likes Mary". If this is an ambiguity, then the theory is inadequate, since it predicts only the first case, which is the backward shifted reading. The second case, which is called the "simultaneous" reading, is not predicted. Specifically, (9) has only one LF, (10a), whose truth condition with respect to a world w, a time t and an assignment g is (10b).

(10) a. PAST [John say [PAST [Bill like Mary]]]
b. There is a time t' before t such that for all worlds w' compatible with what John says at t' in w, there is a time t'' before t' such that Bill likes Mary at t'' in w'

Gennari (2003) proposes a solution to this problem, which basically says that there is no problem. She claims that the truth condition in (10b) is correct for (10a), and that the ambiguity in question is really vagueness, which is due to a property of stative verbs such as *like* which she calls the super-interval property. Stativeness is defined as follows.[4]

(11) A sentence Q is stative iff it follows from the truth of Q at an interval t that Q is true at all instants within t

Thus, if Bill likes Mary at some interval t, it is necessary that he likes Mary at some subinterval t' of t. It follows, then, that if Bill likes Mary at an interval t, it is possible that Bill likes Mary at t' which is a super-interval of t. Applying to the case in (10), this means that if Bill likes Mary at an interval t before John's speech – which is what (10b) says – then it is not excluded that he likes Mary at t' which is a super-interval of t. If we take t'

[4] This definition is taken from Kusumoto (2005).

to include the time of John's speech, we get the simultaneous reading. If t' is understood as non-overlapping with John's speech, we get the backward-shifted reading. Call this analysis the vagueness analysis.

Let us recap. We have introduced the basic idea of the Priorian theory of tense. We have seen that this theory makes correct predictions in some cases, and problematic ones in some others. The latter includes the later-than-matrix reading for relative clauses and the simultaneous reading for some embedded sentences with past tense morphology. We have reviewed solutions to these problems: the scope analysis, and the vagueness analysis. Let us now turn to the discussion of another theory of tense, that of Kusumoto (2005).

In this work, Kusumoto advances some arguments against the scope and the vagueness analysis. First, she points out that the former leads to scope paradoxes. Consider the following two sentences.[5]

(12) a. John introduced no student to any professor who (later) got tenure
b. John introduced each student to a professor who (later) got tenure

In (12a), the indirect object is an NPI. Thus it must be in the scope of the downward-entailing quantifier *no student*. Assuming that the highest position to which *no student* can QR is [Spec,V] (Johnson and Tomioka 1997, Fox 2000), it follows that if the NPI is to be licensed, it must also be inside the scope of the matrix past tense. The Priorian theory predicts that the later-than-matrix is not possible in this case. But it is. As for (12b), surface scope interpretation can go hand in hand with the later-than-matrix reading of the relative clause. In the Priorian theory, this means that the indirect object must scope below the direct object and above the matrix tense. Given standard assumptions, this is a paradox.

The argument Kusumoto makes against the vagueness analysis is basically that Gennari's generalization is wrong. I quote Kusumoto: "[...]when *Karen dance* is true at t, it is possibly true at its super-interval t'. Yet sentences like *Tom said Karen danced* do not allow a simultaneous interpretation." There is reason to think that this argument is not convincing. We come back to this below.

Kusumoto then proposes a theory which purports to account for these facts. First, an additional basic type i is introduced into the set of types, with D_i being the set of time intervals. Second, the following assumptions are made (Kusumoto 2005: 339).

---

[5] Kusumoto's examples are different. She uses control sentences with matrix negative predicates. Here are two of her examples.

(i) a. I tried not to hire anybody who put on a terrible performance
b. She failed to talk to any prospective student who (later) decided to come to UMass.

There are problems with these sentences, as pointed out by Ezra Keshet (p.c.). First, many speakers find (ia) deviant, and agree that the object NP should be a partitive construction such as *any of the actors who put on a terrible performance*. But then we have the possibility of raising the partitive PP, leaving the NPI behind. The second problem is that for (ib), a case can be made that although the talking might precede the decision, the failure must be understood as taking place after the decision (Keshet 2007).

(13) a. Predicates have an extra argument slot for time.
b. Tense morphemes are time variables that saturate the time argument slots of predicates. This means that tense morphemes themselves do not contribute to the meaning of anteriority or simultaneity.[6]
c. The meaning of anteriority and simultaneity derive from phonologically null elements that stand in a certain relation with tense morphemes. These elements give the ordering between eventuality times and evaluation times.

Thus time is no longer an evaluation parameter. It is now represented in the syntactic tree. Let us use a concrete example to illustrate how this theory works. Consider the sentence *John danced*. According to Kusumoto, this sentence has the LF in (14). Denotations of the terms in (14) are given in (15).[7] Calculation of the truth condtion for (14) relative to a world w and an assignment g yields (16).

(14) [t* [PAST [λt_1 [$past_1$ [John dance]]]]]

(15) a. $[\![\text{dance}]\!]^{w,g} = \lambda x \in D_e.\ \lambda t \in D_i.$ x dances at t in w
b. $[\![\text{past}_1]\!]^{w,g} = g(1)$
c. $[\![\text{PAST}]\!]^{w,g} = \lambda P \in D_{<i,t>}.\ \lambda t \in D_i.\ \exists t'$ before t such that $P(t') = 1$
d. $[\![t^*]\!]^{w,g} = s^*$, the speech time

(16) There is a time t' before s* such that John dances at t' in w

In this theory, the relationship between morpho-syntax and semantics is indirect. Specifically, past tense morphology is the result of *past*, whereas past tense interpretation is the result of *PAST*, which is phonologically empty. This opens the possibility of sentences showing past tense morphology but having no past tense interpretation. Thus, *John danced* could have the LF in (17a) and the truth condition in (17b).

(17) a. [$past_1$ [John dance]]
b. John dances at g(1) in w

To rule out this option, Kusumoto follows Stowell and assumes that *past* is like a polarity item. It must be c-commanded by a licensor. What licenses *past* is, unsurprisingly, *PAST*. Thus, (17a) is ill-formed. Note, also, that Kusumoto's system is entirely type-driven. In other word, abstractors, *PAST*, and *t** can be freely inserted up to meaningfulness and convergence.

Let us now see how Kusumoto accounts for the later-than-matrix reading in relative clauses. Let us go back to sentence (5), repeated here in (18a). The LF of this sentence is (18b).

---

[6] Predicates include both nouns and verbs, but only verbs have their time slot saturated in the syntax. Nouns, according to Kusumoto, are tenseless predicates. For present purposes, I consider nouns as predicates with no time slot (see note 2). Also, present tense will not be discussed here.
[7] Kusumoto assumes that worlds are also represented syntactically. But this is not crucial for any point to be made here. I will continue to assume that worlds are evaluation parameters to keep the Priorian theory and Kusumoto's minimally different.

(18) a. Hillary married the man who became president

b. [t* [PAST λt_1 [$past_1$ [Hillary marry [the man [who_2 t* PAST λt_3 $past_3$ t_2 become president]]]]]]

Calculation of the truth conditions for (18b) with respect to a world w and an assignment g will yield (19).

(19) There is a time t' before s* such that Hillary marries at t' in w the unique individual x such that x is a man in w & there is a time t'' before s* such that x becomes president at t'' in w

This truth condition allows for the later-than-matrix reading: it requires both the marriage and the presidency to precede the speech time s*, but it does not impose any temporal order upon the presidency and the marriage themselves.

As for the simulataneous reading, it is derived as follows. Consider again sentence (9), repeated here in (20a). The denotation of the intensional verb *say* is given in (20b).

(20) a. John said Bill liked Mary

b. $[\![\text{say}]\!]^{w,g} = \lambda P \in D_{<s,<i,t>>}. \lambda x \in D_e. \lambda t \in D_i. \forall w'$ compatible with what x says in w at t, P(w')(t) = 1.

Sentence (20a) may have two LFs in this theory, (21a) and (21b). Given Intentional Functional Application (IFA), both are well-formed in the sense that neither has type mismatch and both have every instance of *past* licensed by a c-commanding *PAST*.

(21) a. [t* PAST λt_1 $past_1$ John say [PAST λt_2 $past_2$ Bill like Mary]]

b. [t* PAST λt_1 $past_1$ John say [λt_2 $past_2$ Bill like Mary]]

Calculation of the truth conditions for (21a) and (21b) relative to a world w and an assignment g will yield (22a) and (22b), respectively.

(22) a. There is a time t' before s* such that for all worlds w' compatible with what John says at t' in w, there is a time t'' before t' s.t. Bill likes Mary at t'' in w'

b. There is a time t' before s* such that for all worlds w' compatible with what John says at t' in w, Bill likes Mary at t' in w'

The reading in (22a) is the backward shifted reading and (22b) is the simultaneous reading. Thus, it is possible in this theory to capture the ambiguity of (19). The key to this possibility is the type-theoretic property of *PAST*. It takes an argument of type <i,t> and returns a denotation of the same type. In a type-driven system, this means that the presence of *PAST* is optional, provided its absence does not cause violation of any independent principle, specifically, the licensing condition on *past*.[8]

---

[8] Note that we cannot have the following LF for (20a).

(i) [t* PAST λt_1 $past_1$ John say [t* λt_2 $past_2$ Bill like Mary]

This LF has type mismatch. The verb *say* takes an internal argument of type <s,<i,t>>, but what we have in (i) is something of type <s,t>. Given that the later-than-matrix reading for an embedded past tense sentence requires the presence of t* in that sentence, this explains why (19) cannot have the later-than-matrix reading.

We have seen how Kusumoto's system works. In particular, we have seen how it accounts for the later-than-matrix reading in relative clauses and the simultaneous reading in complement clauses of intensional verbs. If I understand the theory correctly, it leaves some open questions. In what follows, we will discuss these.

First, consider again (17a), which is repeated here in (23a). It seems that nothing in Kusumoto's theory prevents this sentence from having the LF in (23b).

(23) a. Hillary married the man who became president
b. [t* PAST λt_1 $past_1$ [Hillary marry [the man [who_2 $past_3$ t_2 become president]]]]

Evaluation of (23b) with respect to a world w and assignment g will yield the following truth condition.

(24) There is a time t' before s* such that Hillary marries at t' in w the unique individual x such that and x becomes president at g(3) in w

Taken at face value, this is incorrect. Hence we need to impose some restriction on g in cases such as this so that it does not map the relevant variable to some time interval after s*, for example. It is not obvious how this could be done. Another option might be to say that not only must *past* be licensed by *PAST*, it must also be abstracted over by a λ-operator, whereby the only term which can be combined with this λ-abstract is the indexical t*. While this seems straightforward, it also looks rather stipulative.

The second question that Kusumoto (2005) raises is how to rule out simultaneous reading for such cases as (3a), repeated here in (25).

(25) John said Bill kissed Mary.

As far as I can see, nothing in the theory prevents (25) from having the LF in (26a), whose truth condition relative to a world w and an assignment g is (26b).

(26) a. [t* PAST λt_1 $past_1$ John say [λt_2 $past_2$ Bill kiss Mary]]
b. There is a time t' before s* such that for all w' compatible with what John says at t' in w, Bill kisses Mary at t' in w'

Again, this in incorrect. (26b) says that Bill is reported to be kissing Mary at the time of John's speech, which is not in accordance with our intuition. Thus, Kusumoto will have to say that with verbs like *kiss*, *PAST* has to be present in the embedded clause, whereas with verbs such as *like*, it does not have to be. This is basically saying that the availability of the simultaneous reading depends on some aspect of the meaning of the verb, which is not really different from what Gennari says.

Recall Kusumoto's criticism of Gennari's theory: "[…]when *Karen dance* is true at t, it is possibly true at its super-interval t'. Yet sentences like *Tom said Karen danced* do not allow a simultaneous interpretation." I am not sure if this is a valid criticism. It seems to me that when Kusumoto says "sentences like *Tom said Karen danced* do not allow a simultaneous interpretation", she understands *dance* as an achievement verb, i.e. one without the superinterval property, and when she says "when *Karen dance* is true at t, it is possibly true at its super-interval t'," she has in mind the activity version of *dance*, i.e. one that does have the superinterval property. If we construct an embedded context which makes it clear that *dance* is understood as an activity verb, the simultaneous reading seems to be possible. For example, let us say that John called me yesterday, at which time Mary was dancing besides him, and John told me this during our conversation. Given that, I think I can say (27) afterwards.

(27) John said Mary danced the whole time we were talking

Hence it seems possible that Gennari's account, i.e. the vagueness analysis, is really the correct account for the simultaneous reading. In other word, it seems true that the simulataneous reading is available only if the embedded verb has the superinterval property. It might be that Gennari's use of the term 'stative' is unfortunate, because the definition of stativeness actually includes activity verbs which we do not normally call stative, but if that is the case, then there is no issue beyond terminology.

Let us recap. We have reviewed two theories of tense, the Priorian theory and Kusumoto (2005). The set of data against which we evaluate these theories consists, in essence, of the following four sentences.

(28) a. John danced
b. John said Bill kissed Mary
c. John said Bill liked Mary
d. Hillary married the man who became president

For the simple (28a), both theories generate the right reading. For (28b), the Priorian theory predicts only the backward shifted reading, while Kusumoto's theory allows both the backward shifted and the simultaneous reading. Hence, the Priorian theory is better in this case. As for (28c), evaluation of the theories depends on whether Gennari is right. If she is, then both theories are empirically adequate, with Kusumoto's being redundant. If Gennari is wrong, then Kusumoto's theory fares better. For (28d), both theories are inadequate. The Priorian theory undergenerates: it cannot predict the later-than-matrix reading. Kusumoto's theory, on the other hand, overgenerates: it predicts not only the later-than-matrix reading, but also a reading in which the tense of the relative clause is indexical, i.e. fixed by the assignment function and thus does not have to be anterior at all.

Thus, it turns out that if Gennari is right, as I think she is, then the Priorian theory is definitely better than Kusumoto (2005). The only problem faced by the Priorian theory is that it undergenerates in the case of (28d). Specifically, it cannot give us the later-than-matrix reading for relative clauses. On the other hand, Kusumoto's theory overgenerates excessively. The reason, I think, is clear: it postulates too much stuff that can be freely inserted or suppressed. As a concrete example, take a simple sentence like *John danced*. In Kusumoto's theory, the LF of this sentence has two tense pronouns, one bound (*past*) and one free (*t**), one logical constant (*PAST*), and one binder, all represented in the syntax, and all there to implement the temporal shift to the past. Moreover, the only element that ever has morpho-syntactic realization is *past*, which itself has nothing to do with past tense interpretation. Of course, one can think of all kinds of conditions that could be imposed on the distribution of these elements so that the theory would not overgenerate. It is not obvious that this can be done in a non ad hoc manner, if at all. Let us, then, assume that the Priorian theory is closer to the truth.

At this point, the obvious question is: what about the later-than-matrix reading? I have no real answer to this question, and the paper could end here. However, I will venture a suggestion as to how an answer could look like. I hope to be able to work out a concrete proposal in the future.

Let us start by adding some more observations into the data set. It seems that the distinction between the backward shifted and the later-than-matrix reading appears in more cases than we have considered so far. Until now, we have looked at past tense interpretation of CPs that fall under one of the following two categories.

(29) a. Complements of VP (John said Bill kissed Mary)
b. Adjuncts of NP (Hillary married the man who became president)

The generalization is that VP-complements show backward shifted, while NP- adjuncts show later-than-matrix reading. Let us, now, consider the two remaining possibilities: VP-adjuncts and NP-complements.

(30) a. John died before/if Bill kissed Mary (VP-adjunct)
b. John presented the evidence that Mary went to NY (NP-complement)

It is evident that VP-adjuncts show later-than-matrix while NP-complements show backward shifted reading. The generalization, then, is (31). For presentational clarity, the representative data are gathered into a paradigm in (32) and (33).

(31) In a past under past construction, the later-than-matrix reading is possible only if the embedded past tense is inside an adjunct

(32) Complements & backward shifted reading
a. John said [that Bill kissed Mary]
b. John presented the evidence [that Mary went to NY]

(33) Adjuncts & later-than-matrix reading
a. John died [before Bill kissed Mary]
b. Hillary married the man [who became president]

What follows will be an attempt to make sense of (31). First, note that tense interpretation is not the only phenomenon in which we see a distinction between complements and adjuncts. Another one which has become famous is the so called Lebeaux effect. Consider (34a-b).

(34) a. * [which claim that John$_1$ cheated]$_2$ did he$_1$ deny t$_2$
b. [which claim that John$_1$ made]$_2$ did he$_1$ deny t$_2$

The Lebeaux effect receives an elegant minimalist analysis, and it is this analysis that I will take as the starting point of my tentative proposal for deriving (31). Let us, then, briefly go through the explanation. It must be noted, however, that what I present here is a particular version of the account, extremely simplified to fit the purpose and scope of this squib.

Suppose that a sentence violates Condition C if at some point in its derivation, an r-expression is c-commanded by a coindexed DP. This straightforwardly accounts for the ungrammaticality of (34a), since at one point in its construction, the derivational workspace contains the following syntactic object.
(35) [he$_1$ deny [which claim [that John$_1$ cheated]]]

But what about (34b)? Here is the key assumption that explains the contrast between (34a) and (34b).

(36) Late Adjunction Hypothesis (LAH)[9]
Adjuncts can merge late, i.e. after the main cyclic derivation has been completed, whereas complements must merge with the selecting head before, that head merges with any other constituent.

The LAH allows the CP-adjunct in (34b) to merge after wh-movement has taken place. Thus, at one point in the derivation of (34b), the derivational workspace contains two syntactic objects, namely those in (37).

(37) a. α = [$_{CP}$ [which claim]$_2$ did he$_1$ deny t$_2$]
b. β = [$_{CP}$ OP$_1$ that John made t$_1$]

The next step is the late merger of β into α, resulting in (38).

(38) [$_{CP}$ [which claim [$_{CP}$ OP$_1$ that John made t$_1$]]$_2$ did he deny t$_2$]

At no point in the derivation is Condition C violated. (34b) is predicted to be good. Thus, the Lebeaux effect is derived.

It would be very hard, if possible at all, to explain the Lebeaux effect in GB terms, since the crucial ingredient to the explanation is the radical derivationalism of syntactic structure building which distinguishes minimalist syntax from its immediate predecessor. In GB, an initial phrase marker is projected all at once by X'-Theory and θ-Theory, which is then manipulated by Move α. In minimalism, on the other hand, syntactic structure is constructed step by step, with merge and move interspersing throughout the derivation. And recently, it has been proposed that not only the construction but also the interpretation of syntactic structure proceeds step by step. For example, Chomsky (2006) suggest that PF and LF be dispensed with. What remains is an operation, Transfer, which delivers pieces of syntax to the interpretive components SM and CI.

This way of thinking sets the stage for a natural explanation of cyclicity. The general form of the explanation is that pieces which have been sent off to the interpretive components become inaccessible. Questions of details then arise. What are the properties of these pieces? In what sense are they inaccessible? Are the pieces that get sent to SM different from the pieces that get sent to CI. What happens at SM and CI? Finding answers to these specific questions is essential to the articulation of a minimalist theory.[10]

Coming back to our problem, i.e. the later-than-matrix reading in adjunct CPs, the tentative solution that I am going to propose for this problem will analyze it as basically a case of cyclicity. If it is on the right track, it might constitute a step toward clarifying the nature of the interaction between narrow syntax and CI.[11]

---

[9] If θ-roles are formal features that have to be checked (Fanselow (2001)), the fact that complements have to merge early will follow from Featural Cyclicity (Chomsky (1995), Richards (1998)).
[10] For a concrete proposal concerning the sound side, see Fox and Pesetsky (2005).
[11] If I understand Chomsky correctly, my use of the term "CI" is different from his, as will be clear below. It might be better to say "the semantic system" instead, since what is meant here seems much more linguistic than what Chomsky intends with "CI".

Let us first consider CI itself. Essentially, it is a system which pairs structures with meanings. Let us, then, assume just that: CI takes transferred SOs and forms a set of pairs of the form $<\alpha, \alpha'>$, where α is an SO and α' its interpretation. For present purposes, let us assume that the interpretation at CI is model theoretic. Specifically, CI will supply a specified index and an arbitrary assignment with respect to which denotations of syntactic objects are to be computed.[12]

For illustration, let us consider an abstract example. Suppose that at some stage in the derivation of a sentence S, two SO's, say α and β, are lying around in the derivational workspace. Suppose that conditions for Transfer are met. Transfer will then apply, sending α and β off to CI. At CI, α and β will be evaluated with respect to a certain index, say (w,t), and an arbitrary assignment, say g. The set D is then formed.

(39) $D = \{<\alpha, \llbracket\alpha\rrbracket^{w,t,g}>,<\beta, \llbracket\beta\rrbracket^{w,t,g}>\}$

Call D the denotation set. Now suppose that the syntactic derivation continues, merging β with α, forming γ, and suppose that conditions for Transfer are again satisfied. Then γ will be sent off to CI. The question is what happens at CI now. It seems implausible that CI will just add $<\gamma, \llbracket\gamma\rrbracket^{w,t,g}>$ into the denotation set. The reason is this. If with each Transfer, a new pair is simply added into the denotation set obtained in the last Transfer, there might come a point where the syntactic derivation terminates and we end up at CI with a set of disparate structure-meaning pairs. Intuitively, this is not what a sentence is. A sentence is one piece of structure paired with one piece of meaning. Hence, let us say that each time Transfer applies, the denotation set is updated. Specifically, the new denotation set will consist of the most recently transferred SO paired with its interpretation. In our example, D in (37) will be updated to D' in (40).

(40) $D' = \{<\gamma, \llbracket\gamma\rrbracket^{w,t,g}>\}$

If we assume, following Chomsky (1995), that a derivation converges only if the Numeration is empty and the workspace contains exactly one object, then it is ensured that the final denotation set will consist of exactly one member.

But if this is all there is to it, we might just as well say that only the final product of the syntactic derivation can be transferred, and there will be no cyclicity effects. So of course this is not all there is to it. The computation of $\llbracket\gamma\rrbracket^{w,t,g}$, i.e. the newly transferred object, is subject to a condition. It is this: for every portion π of γ, if π is identical to an SO in the previous Transfer, then π must receive the interpretation given to that SO in the previous Transfer. In short, once an SO receives an interpretation M, it will always be paired with M in further computation.[13]
Now let us consider Transfer. Let us ask what the conditions are that have to be met in order for Transfer to apply? For CI, it is reasonable to assume that the transferred objects are categories that typically correspond to complete sentences. Let us, then, assume the following.

---

[12] An index is a pair (w,t) of world and time. By "an arbitrary assignment", I want to say that the assignment function can vary through computational steps. This will allow variable binding into adjuncts,
[13] We have to assume, of course, that CI can look at and compare syntactic objects. This is actually not an implausible assumption. In fact, it might just follow from the nature of CI: a system that pairs syntactic structures with their denotation. As for the problem of determining identity, this must be taken care of anyway, independently of any assumption about CI. Chomsky (1995) proposes a way to do this with the Numeration. I will assume here that the problem has a solution.

(41) A syntactic object SO is transferred to CI iff the following holds:[14]

a. SO is a CP
b. SO has no uninterpretable features
c. the Numeration is empty

This will get us the later than matrix reading in adjunct CPs. Here is how. Consider again our famous sentence.

(42) Hillary married the man who became president

According the LAH, we can choose to insert the adjunct late. If this option is taken, then at some point in the derivation of (42), the workspace will contain the following syntactic objects.

(43) a. α = [$_{CP}$ Hillary married [the man]]
b. β = [$_{CP}$ who became president]

As the Numeration is empty, α and β are transferred to CI. The denotation set in (43) is formed.[15]

(44) $D = \{<\alpha, [\![\alpha]\!]^{w,t,g}>,<\beta, [\![\beta]\!]^{w,t,g}>\}$

Next, the syntactic derivation continues, inserting β into α, forming γ as in (43).

(45) γ = [$_{CP}$ Hillary married [the man β]]

Transfer will again apply, sending off γ to CI. The denotation set in (44) is formed.

(46) $D' = \{<\gamma, [\![\gamma]\!]^{w,t,g}>\}$

Given what we assume, CI will scan γ, and if it sees any portion π of γ that is identical to α or β, it will assign to π the same denotation it assigns to α or β in the previous Transfer, and use that denotation in its computation of aγ b w,t,g. In this case, there is such a portion: the relative clause. While α has been extended to α', which differs from α in that it contains β, β itself remains intact. Thus, the contribution which β makes to the calculation of $[\![\gamma]\!]^{w,t,g}$ will be $[\![\beta]\!]^{w,t,g}$. Given that, it turns out that $[\![\gamma]\!]^{w,t,g} = 1$ iff (47) hold.

(47) There is a time t' before t such that Hillary marries at t' in w the unique individual x such that there is a time t'' before t such that x is a man in w and x becomes president at t'' in w

---

[14] Danny Fox suggested that while phonology can interpret the sentence piece meal, semantics is likely to take the whole sentence as input (Spring 2007 MIT class lecture). Although what I propose here is not exactly what Danny said, it is in the same spirit.

[15] Technically, the denotation of α will be a truth value in this case, which does not really accord with our intuition about the meaning of a sentence. However, we can say that knowledge of meaning is really knowledge of the computation of denotations. In that case, we would have to know under which conditions the sentence gets the value 1, which is what the meaning of the sentence is.

This is the later-than-matrix reading. The way it comes about reminds us of other cyclicity effects. In particular, the *PAST* operator inside β is sealed off from the matrix *PAST* operator, i.e. the index with respect to which β is evaluated is immune to the shifting effect of the matrix *PAST*. The result is just what we want: as far as evaluation time is concerned, β behaves as if it is not embedded.

The proposal here also accounts for the fact that complement CPs always show backward shifted reading. Let us show this by way of another familiar sentence.

(48) John said Bill kissed Mary

At some point in the derivation of (48), the workspace contain the following object.

(49) α = [$_{CP}$ Bill kissed Mary]

If Transfer can apply to α as such, we would predict that the later-than-matrix reading is available for the embedded CP in (48), which is not the case. It turns out that given the system proposed here, Transfer cannot apply to α as such. Here is why. In order for Transfer to apply, the Numeration must be empty. Now it is reasonable to assume that lexical items (LIs) are taken out of the Numeration one by one, and when an LI is taken out of the Numeration, it must undergo merge as soon as it can. Given that, the only situation in which α can undergo Transfer is one in which the derivational workspace contains two syntactic objects, namely the following.

(50) a. α = [$_{CP}$ Bill kissed Mary]
b. β = [$_{CP}$ John said]

But this would mean that *say* merges with *John* before it merges with α, a violation of the LAH, which dictates that a head must merge with its complement before it merges with any other XP.

Thus, the only object that can undergo Transfer in this case is γ in (51).

(51) γ = [$_{CP}$ John said [$_{CP}$ Bill kissed Mary]]

This means that when the embedded *PAST* gets to CI for the first and only time, it is already embedded under the matrix *PAST*. Thus, the only denotation set in this case will be (52).

(52) D = {<γ, $[\![\gamma]\!]^{w,t,g}$>}

Calculation of $[\![\gamma]\!]^{w,t,g}$ will yield 1 iff the the condition in (53) holds. This is the backward shifted reading.

(53) There is a time t' before t such that for all w' compatible with what John said in w, there is a time t'' before t' such that Bill kisses Mary at t'' in w'

Basically, the later-than-matrix reading obtains when two CPs, the matrix and the embedded, behave as if neither contains the other. This is possible only if there is some point in the derivation at which the two CPs are transferred to CI, and at which it is actually the case that neither CP contains the other. This can never happen if the embedded CP is a complement, but it can happen if it is an adjunct. Hence, only embedded CPs that are adjuncts show later-than-matrix reading.

We have basically solved the problem we set out to solve. Now the theory proposed here has some empirical predictions that go beyond the data we have considered. I will now discuss these. Let us start with sentence (54).

(54) John wants to find the woman who is a man

Imagine the following situation. John met a woman at the post office, fell in love, but kept quiet. Now he wants to find this woman to tell her that he loves her, but she is nowhere to be found. I, the speaker, know that the woman who is occupying John's mind is really Bill, who dressed up as a woman on the day John went to the post office. I think I can say (54) in this case. Thus, there is a reading of (54) such that it is true relative to a world w, a time t and and assignment g iff the following holds.

(55) For all w' compatible with what John wants at t in w, John finds the (contextually) unique x such that x is a woman in w' and x is a man in w

It turns out that there is a straightforward way to derive this reading in our theory. Suppose that we decide to late merge the relative clause. At some point in the derivation, we have two SOs lying around in the workspace.

(56) a. α = [$_{CP}$ John wants to find the woman]
b. β = [$_{CP}$ who$_1$ t_1 is a man]

Transfer applies, sending α and β to CI. At CI, these objects are evaluated with respect to an index (w,t) and an assignment g. Calculation of $[\![\beta]\!]^{w,t,g}$ will yield (57) as the result.

(57) $[\![\beta]\!]^{w,t,g} = \lambda x.$ x is a man in w at t

It is this denotation that will be utilized in the calculation of the final syntactic product, call it γ, which is constructed by inserting β into α. Computation will yield the truth condition in (55) for γ.

Let us discuss another prediction of the theory. Consider the following sentence.

(58) John looked for [$_{\beta}$ the evidence [$_{\alpha}$ that Mary went to New York]]

By hypothesis, α is the complement of [$_{N}$ *evidence*]. Given what we have said, there can be no later-than-matrix reading. In other word, the time at which Mary went to New York must preceed John's search attempt. However, the theory also predicts that if for some reason, β can raise above the matrix T, then the later-than-matrix reading will be possible. Moreover, this reading is possible <u>only if</u> β raises above the matrix T.

Since T is obviously above V, we expect that if α is to receive the later-than-matrix reading, β must scope over the intensional verb. In other word, we expect the later-than-matrix reading to correlate with a *de re* interpretation of β. This seems to be right. Thus, in order to get the later-than-matrix reading for α in (58), we have to imagine such a situation as follows. We are investigating a crime that took place in New York. The suspect is Mary, who lives in Chicago. At the present, we know that Mary went to New York last Wednesday. In addition, we know that two weeks ago, Mary bought a plane ticket for $300 which bears her name as passenger and New York as the destination.

This plane ticket is thus the evidence that Mary went to New York. We also know that John, who is Mary's husband, found out last Tuesday, i.e. the day before Mary went to New York, that $300 are missing from his safe, went nut, rummaged his wife's closet in order to find out what she had bought. I think that we can say (56) in this case. The sentence sounds better if we change it a little bit, as in (57).

(59) Now we know that on Tuesday, John looked for the evidence that Mary went to New York

Above, we said that [Spec,V] is the highest position that an object can raise to. Obviously, this is not totally right. However, to spell out the conditions under which object NPs can raise above T is beyond the scope of this squib. It suffice to say that what is predicted by the theory is born out by fact: in such sentences as (57), the later-than-matrix reading must go together with a *de re* interpretation of the object, and a *de dicto* reading must correlate with a backward shifted reading.

Let us conclude. We have presented the pros and cons of two theories of tense interpretation: the Priorian theory and Kusumoto (2005). We have argued that the Priorian is empirically more adequate. However, that theory as such is not able to predict the later-than-matrix reading in relative clauses. We have proposed a tentative solution to this problem. The proposal turns out to predict some additional facts beyond the data it intends to cover.

References

Chomsky, N. (2006). Approaching UG from below. Ms: MIT.

Enç, M. (1987). Anchoring Conditions for Tense. *Linguistic Inquiry* 18: 633-657. Fanselow, G. (2001). Features, theta-roles, and free constituent order. *Linguistic Inquiry* 32: 405–437.

Fox, D. (2000). *Economy and Semantic Interpretation*. MIT Press.

Fox, D. and Pesetsky, D. (2005). Cyclic linearization of syntactic structure. *Theoretical Linguistics 31*:1 - 45.

Gennari, S. P. (2003). Tense Meanings and Temporal Interpretation. *Journal of Semantics* 20: 35.

Johnson, K. and Tomioka, S. (1997). Lowering and mid-size clauses. *Proceedings of the 1997 Tübingen Workshop on Reconstruction*: 185 - 206.

Keshet, E. (2007). Handout on Kusumoto (2005), presented at MIT's LF Reading Group on March 8, 2007.

Kusumoto, K. (2005). On the Quantification over Times in Natural Language. *Natural Language Semantics* 13: 317-357.

Ladusaw, W. (1977). Some Problems with Tense in PTQ. *Texas Linguistic Forum* 6: 89-102.

Montague, R. (1973). The proper treatment of quantification in ordinary English. *Approaches to Natural Language* 49: 221–242.

Prior, A. N. (1967). *Past, Present and Future*. Oxford University Press.

Richards, N. (1998). The principle of minimal compliance. *Linguistic Inquiry* 29: 599 - 629.

Stowell, T. (1993). Syntax of Tense. *Ms., UCLA*.