August 5, 2007

Head Movement, Binding Theory, and Phrase Structure

Junri Shimada

Introduction

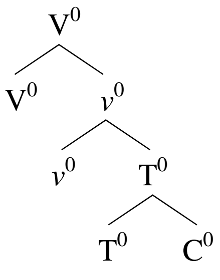
This paper proposes a novel model of head movement that solves two seemingly independent problems.

The first problem concerns the syntactic characteristics of head movement. Head movement appears to move a head and adjoin it with a non-root node of the syntactic tree. In the minimalist framework, however, only the root node is accessible for syntactic operations and the condition that a moved element should c-command its trace is a direct consequence of this assumption. Thus, head movement poses a serious problem to minimalist syntax.

The second problem concerns the binding constraint discussed by Percus (2000). Percus observes that the world variable of the main predicate, if syntactically present, always has to be bound by the nearest λ operator. In addition, I discuss that we can make a similar generalization regarding Tense as well. However, configurationally, it is conceivable that such a variable might be bound long-distance, and it is not clear why such a binding may not obtain.

To solve these problems, I propose a derivation model in which heads are combined to form a complex head that looks like something like (1) and derivation proceeds as heads move out of the head complex by head movement.

(1)



Thus, merger of heads is done before head movement, and therefore the first problem mentioned above is obviated, since the moved head does not adjoin to another head after head movement and thus the moved head is always going to c-command its trace. The idea is that moving heads are quantifiers and have to move out to remedy a semantic type mismatch just like QR. I develop semantics in which the functional heads denote quantifies over non-individual elements, and given the semantics, we see that Percus's generalization is automatically derived by this model. I explore the appropriate locality condition for head movement in my model, and show that the non-attested long-distance configurations are excluded because that would have to involve non-local head-movement.

The organization of the paper is as follows.

Chapter 1 explains the two above-mentioned problems in detail.

In Chapter 2, I show the outline of my derivation model. I then show how various word orders in different languages can be derived in my model. In the last part of this section, I discuss a possibility of eliminating the head-directionality parameter.

In Chapter 3, I develop semantics that is compatible with my derivation model. Particularly, I consider interpretations of presuppositional and cardinal noun phrases and determine the internal structure of noun phrase. More precisely, I propose that while presuppositional noun phrases have their own quantifications over situations, cardinal noun phrases take a situation variable that is an anaphor.

In Chapter 4, I develop the syntax of my model in detail. I propose a new phrase structure according to which phrases are formed only by virtue of Internal Merge and an axiom that only maximal projections may move. I determine the locality for head movement and derive properties of head movement as theorems. Finally, I show how Percus's generalization follows from my model.

Finally, Conclusion gives the conclusion.

Chapter 1 Problems in Head Movement and Binding Theory

In this chapter, I present seemingly two unrelated problems that I wish to solve in this paper. The first of them concerns the syntactic characteristics of head movement, and the second is the binding constraint discussed by Percus (2000)

1.1. Head movement

Pollock (1989) argues that French verbs are raised past an adverb attached in the VP, while English verbs are not.

(1) a. *John kisses often Mary.

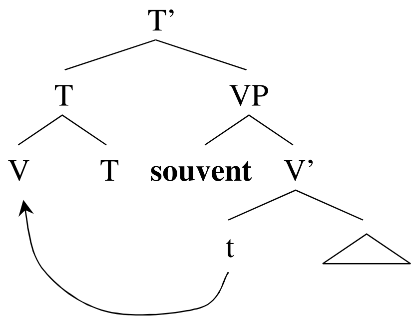
b. Jean embrasse souvent Marie.
John kisses often Mary

c. John often kisses Mary.

d. *Jean souvent embrasse Marie.
John often kisses Mary

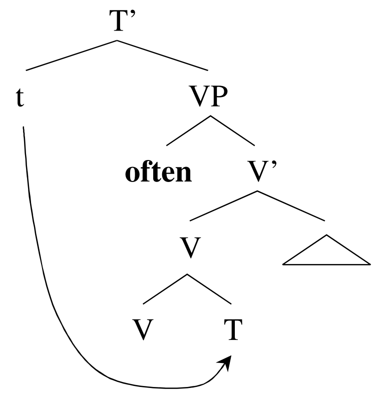
Pollock argues that this contrast is explained by assuming that V is raised to the T position by head movement in French, while normal English verbs like *kiss* do not undergo such movement. This V to T movement has often been analyzed as involving a configuration as follows.

(2) V to T movement



On the other hand, assuming that a tensed verb (e.g. *kiss*) in English contains a bare verb (e.g. *kiss*) and an inflectional ending (e.g. *-s*), T seems to lower onto V in English:

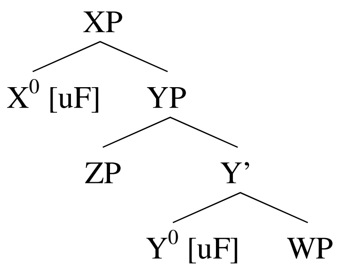
(3) T to V movement



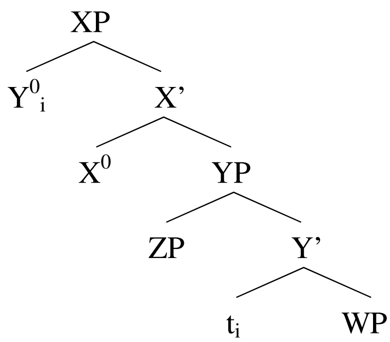
In light of the Minimalist Program framework proposed in Chomsky (1995), however, both of these instances of head movement appear problematic. According to Chomsky's Extension Condition, Merge always takes place at the root of a tree, i.e., Merge should add a sister to an existing tree. Matushansky (2006) notes that the landing site of head movement such as in (1) does not c-command the extraction site, and therefore poses a problem.

In order to solve this, Matushansky proposes that head movement is a 2 step process. First, the uninterpretable features in (4a) cause the head Y^0 to move to the root of the tree as shown in (4b) just like phrasal movement. After this, morphological merger applies to the two adjacent heads, yielding the familiar configuration in (4c).

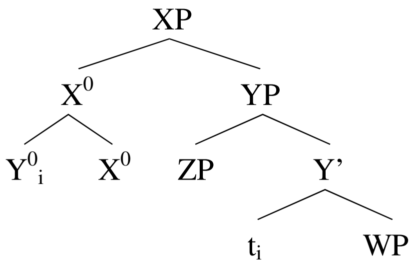
(4) a.



b.



c.



Matushansky defines a head as a syntactically indivisible bundle of formal features, and gaining insight from Distributed Morphology, she proposes that morphological merger is an operation of creating a feature bundle out of the two heads. According to this view, the created bundle is a syntactically single node and so it does c-command the trace of the moved head.

Matushansky's theory is not a solution, however. First, Matushansky's proposal does not seem to solve the very problem that it tries to. In the bare phrase structure proposed in Chomsky (1994, 1995), every syntactic object is either a lexical item picked from the lexicon or a set that is created by recursive application of the binary operation Merge on syntactic objects. Merge yields a non-terminal syntactic object K by merging two syntactic objects α and β:

(5) $K = \{\gamma, \{\alpha, \beta\}\}$

Here, γ is the label of K, and is either the head of α or β (written as $H(\alpha)$ and $H(\beta)$ respectively). In the bare phrase structure, a syntactic tree is merely a way of informally representing a given syntactic object for expository purposes. Now, let us translate Matushansky's proposal into a set theoretic notation in bare phrase structure. Since Chomsky's original notation gets very cumbersome when describing complex syntactic objects, I will adopt an alternative notation such as the folloing.

(6) $K = \gamma{:}\{\alpha, \beta\}$

Although (6) looks different from (5), it should be understood as a mere alternative notation of (5). Adopting this notation, (4b) and (4c) above are considered as informal representations of (7a) and (7b) below respectively:

(7) a. X:{Y, X:{X, YP}}

b. X:{<X,X>:{Y, X}, YP}

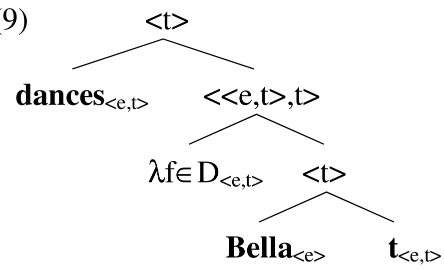
Here, <X,X>:{Y, X} represents a syntactic object that is creating by adjoining Y to X. In order to transform (7a) into (7b), Y in (7a) will have to look inside its sister X:{X, YP}, breaks it down into its two constituents and have itself merged with one of these constituent, namely X, forming <X,X>:{Y, X}. Then <X,X>:{Y, X} and the remaining constituent YP will be merged. However, all these operations are illicit in the minimalist analysis. Given (7a), the only syntactic object we can construct with it is the following:

(8) $\gamma{:}\{\alpha, X{:}\{Y, X{:}\{X, YP\}\}\}$

Thus, although the minimalist approach allows access only to the very root of the tree, morphological merger as proposed by Matushansky actually requires access to non-root nodes in the tree. This is the fundamental problem that previous analysis of head movement have, and although Matushansky claims to offer a solution to it, it really does not.

Second, based on the observation that for example, *John often kisses Mary* and *Jean embrasse souvent Marie* have the same meaning, Matushansky argues that head movement does not have semantic consequences, and it can be explain by assuming the moving heads that are predicates (i.e., type <e,t>) leave a trace of the same type as itself as illustrated below:

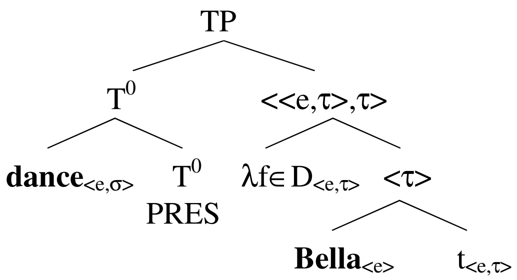
(9)



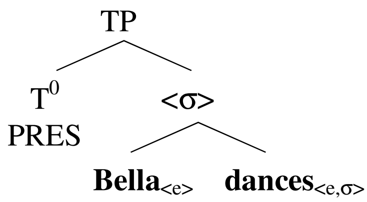
While her generalization that head movement has no semantic effects seems correct, it turns out that her explanation does not quite seem to work, once we take into account the meaning of functional heads such as Tense.

Obviously, Tense is in some way contributing to the sentence meaning as to where in the timeline events etc. expressed in the proposition are located. Since its meaning is transparent (i.e., does not vary depending on which verb it attaches to), we should not have separate semantic lexical entries for *dances* and *danced*. Rather, we should have a single lexical entry for the verb *dance*, and obtain the meaning of different tensed forms as a result of applying the meaning of different Tenses to the tenseless sentence. Thus, we should not disregard the meaning of Tense. With this is mind, consider the two configurations in (10). (10a) represents a case of V to T raising, and (10b) a case where such raising does not take place (we can assume a checking theory account. See the discussion below).

(10) a. V to T raising



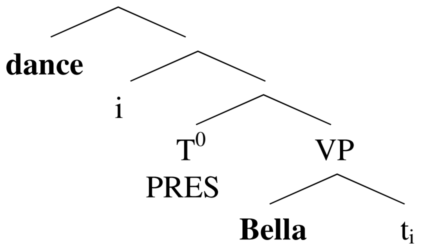
b. No raising



Since head movement does not result in any change in meaning, these two must have the same interpretation. Now, since *Bella* is of type e, *dance* must be of type <e,σ> for some semantic type σ. Likewise, the trace of *dance* must be of type <e,τ> for some semantic type τ. From (10a), we see that PRES is either or type <e> or type <<e,σ>,ρ> for some semantic type ρ. The latter possibility is excluded, however, since then, PRES won't be able to combine with ⟦**Bella dances**⟧$\in D_\sigma$ as it should in (10b). Therefore, PRES must be of type e. However, this is an absurd conclusion given what Tense is. Thus, it is actually untenable to propose that head movement results in various displacements of heads in LF without any semantic effects or problems.

It should be noted that the semantic input in Matushansky's framework should be a configuration where the morphological merger has already taken place as in (10a), and not a configuration like (11) below, where the morphological merger has not taken place:

(11)



If a configuration like (11) were the input to the semantics, the problem discussed above would of course not arise, since as long as the trace left behind by *dance* is of the same type as the denotation of *dance*, the whole structure is interpreted as if there were no movement. In Matushansky's framework, morphological merger of heads is a syntactic process that takes place obligatorily when possible, since the merged head should be able to undergo further successive head movement. Therefore, the moment *dance* is moved by head movement in (11), it will be merged with T^0 immediately, yielding the structure in (10a).

Thirdly, Matushansky cannot treat with what appears to be T to V lowering as a case of true lowering as illustrated in (3). Lowering movement, if is existed, would be truly problematic, since not only would it violate the Extension Condition, but is would also leave the trace of the moved head unbound. Therefore, one would have to resort to the account that assumes that verbs are equipped with their inflection together with some feature to be checked off by T. This way, T does not have to lower. However, this account runs into the same trouble just described above. Travis (1984) argues that head movement is strictly local, and cannot skip an intermediate head (Head Movement Constraint). Assuming that the feature checking is also part of overt head movement such as V to T raising, the feature checking should be strictly local as well. Then, given that there is a *v* head between T and V, in order for T to be able to check off V's features, V must have been already raised to the *v* position. Such movement should be necessary in any case, given the Phase Impenetrability Condition in Chomsky (2001), which states that the complement of *v* is not accessible from outside of *v*P. But then, we would just run into the problem that displacement of heads at LF creates problems for semantics. In sum, neither the lowering analysis nor the checking theory analysis works for so-called T to V lowering.

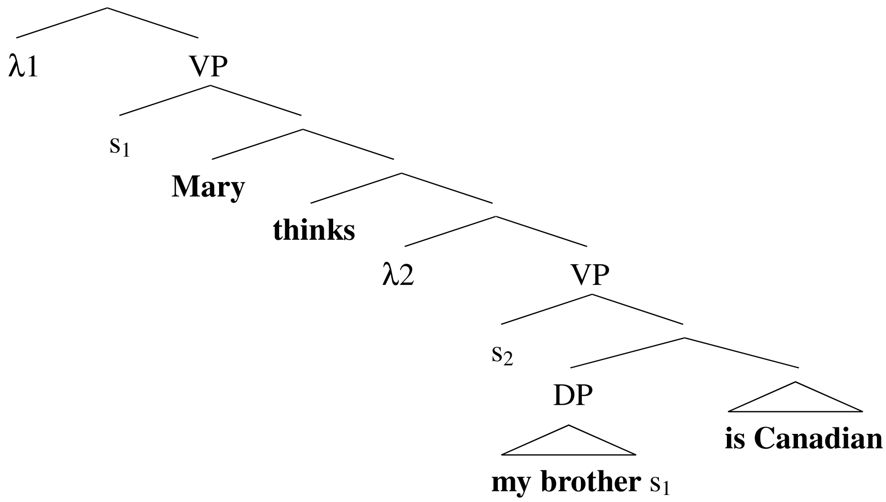
1.2. Percus's problem

Percus (2000) observes that if world variables are syntactically present in syntax, there is a certain constraint that hold on them. Consider the following sentence:

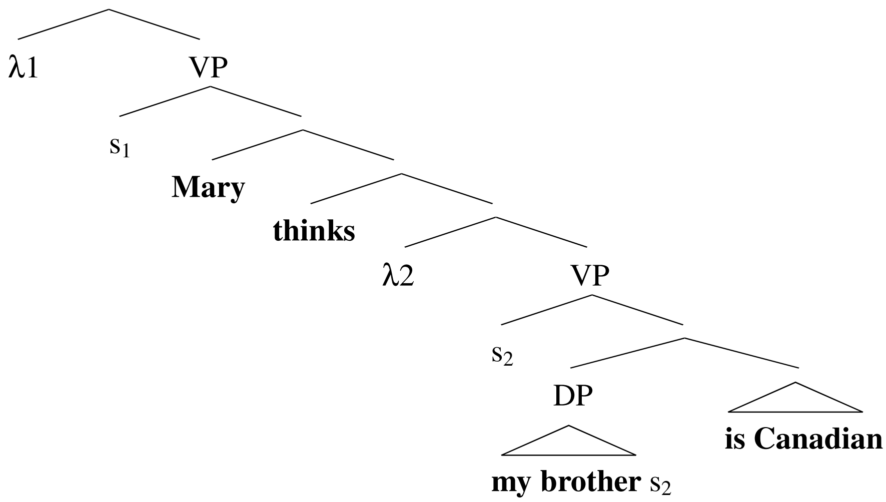
(12) Mary thinks my brother is Canadian.

The sentence has three predicates, namely, *thinks*, *brother*, and *is Canadian*. Each of these predicates has a world argument that is present in syntax. Now if we introduce λ abstractors so as to properly interpret the embedded clause, we have two possible structures for this sentence (the topmost λ abstractor is introduced so that we can apply the intention of the matrix clause to the actual world):

(13) a.



b.



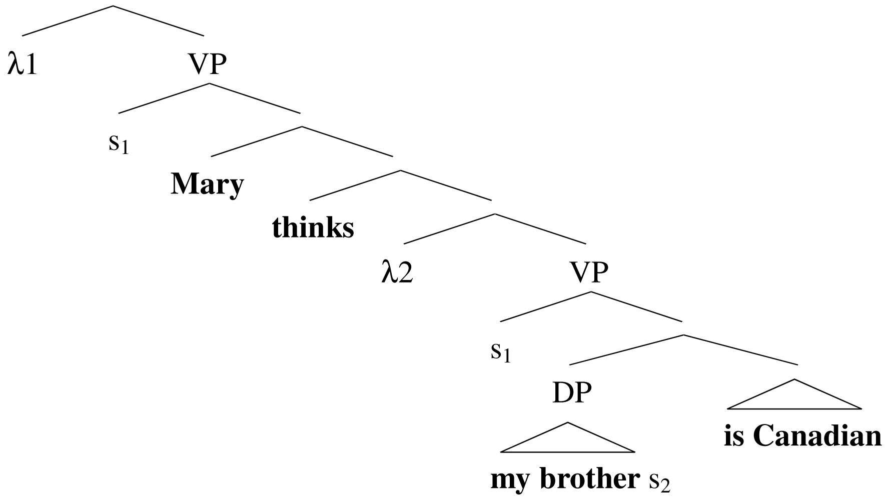
These two would be translated as something like the following.

(14) a. For every world w that is consistent with what Mary thinks in the actual world w*, the unique brother of mine <u>in w*</u> is Canadian in w. (for a)

b. For every world w that is consistent with what Mary thinks in the actual world w*, the unique brother of mine <u>in w</u> is Canadian in w. (for b)

In other words, *my brother* has a *de re* reading in (14a), whereas it has a *de dicto* reading in (14b). Thus, the world variable in the DP *my brother* can be bound by either the higher λ or the closer λ. Percus notes, however, that the binding shown below is impossible.

(15)



Here, the world variable in the DP *my brother* is bound by λ2 and the world variable taken by the predicate *is Canadian* is bound by the higher binder λ1. Although this is a computable structure, the resulting interpretation is something we never get:

(16) For every world w that is consistent with what Mary thinks in the actual world w*, the unique brother of mine in w is Canadian in w*.

This can be paraphrased as "There is an actual Canadian that Mary thinks is my brother". To exclude such a structure, he proposes the following generalization.

(17) Generalization X: The situation pronoun that a verb selects for must be coindexed with the nearest λ above it.

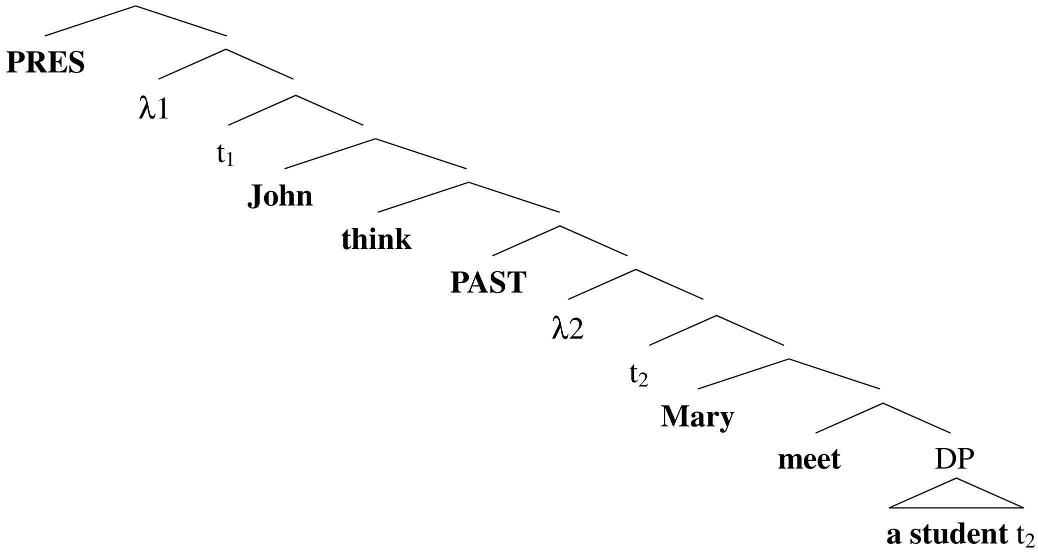
By "situation", he means a part of a world and he actually talks in terms of situations in his paper, but for our purposes here, his "situations" could be regarded as worlds.

What is more interesting about Percus's generalization is that binding theory of this sort does not seem to be limited to worlds (or situations). Kusumoto (2005) and Fox & Heim (2006) who follow Kusumoto suggest the possibility of extending the generalization to time variables. These authors discuss binding of the reference time variable in an embedded clause by a tense operator in the higher clause. Since their discussion assumes a specific model regarding the reference time, I will instead discuss binding of the time variable that the main predicate in an embedded clause takes. Consider the following example.

(18) John thinks Mary met a student.

In this case, we can think of the time variables that predicates take and their binding. Assume operators **PRESENT** and **PAST**, the following structure seems possible, where the time variable in the DP is bound by the closer λ, as it yields a correct interpretation.

(19)

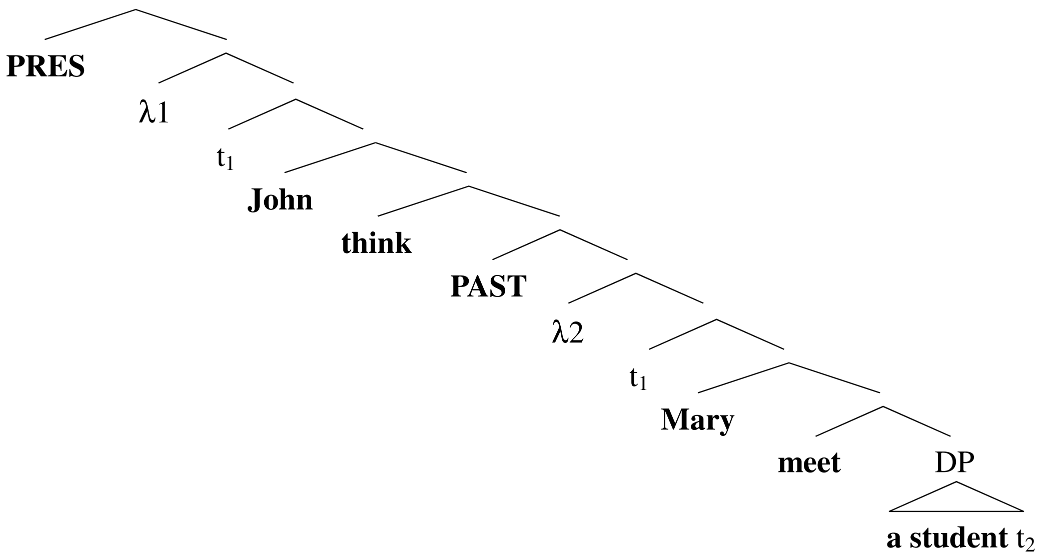


The interpretation of this tree would be something like the following:

(20) In every world w that is consistent with what John thinks now in the actual world, there is a time t such that there is an individual x such that x is a student at t in w and Mary meets x <u>at t</u> in w.

However, the following structure where the time variable taken by *meet* is bound by the higher λ1.

(21)



This would yield the following interpretation, which is clearly unavailable.

(22) In every world w that is consistent with what John thinks now (= t*) in the actual world, there is a time t such that there is an individual x such that x is a student at t in w and Mary meets x <u>at t*</u> in w.

In Heim & Kratzer's (1998) system, a λ abstractor is introduced whenever a movement takes place, and the introduced λ binds the trace of the movement. Percus suggests that perhaps λs that bind world variables are introduced by movement of something like a silent relative pronoun. He notes that his generalization is then explain if such movement out of DPs is somehow prohibited. This is actually the line I am taking, and I propose below that this movement is actually head movement, and that the constraint is the locality constraint of head movement.

1.3 Summary

In this chapter, we have looked at two problems. Although these appear to be rather distinct problems, it turns out that they are connected since head movement creates binding of worlds, times, etc., as I discuss in the rest of this paper.

Chapter 2 Head Complex and Word Order

The main problem that previous analyses of head movement face is that the two (or more) morphologically merging heads originate in positions that are not adjacent to each other. Therefore, such analyses need to have heads morphologically merged as the derivation proceeds. However, that would require operations that are not compatible with the minimalist perspective as we have seen in the previous chapter. In this chapter, I propose a new model of derivation that avoids this problem, and explore how various word orders may be derived in the model.

2.1 Quantification over times – Kusumoto's thoery

Kusumoto (2005) argues that a tense system with a single temporal index like Priorian tense logic (Prior 1967) does not work, because it is not capable of dealing with sequence-of-tense phenomena and what she calls later-than-matrix interpretation. (1a) had two readings paraphrased in (1b) and (1c). The reading in (1b) cannot be explained by a single temporal index system.

(1) a. Tom said that Karen was dancing.

b. Tom said, "Karen is dancing."

c. Tom said, "Karen was dancing."

Also, in addition to the reading that Hilary married a man who had already become the president, (2) also has the reading that Hilary married a man who later became the president. For this latter reading, the relative clause would refer to the utterance time, which option should not be available in a single temporal index system.

(2) Hilary married a man who became the president of the U.S.

In order to solve these problems, based on Stowell (19?), Kusumoto proposes the following model. Predicates have an argument slot for a time (type <i>), as well as a slot for a world (type <s>).

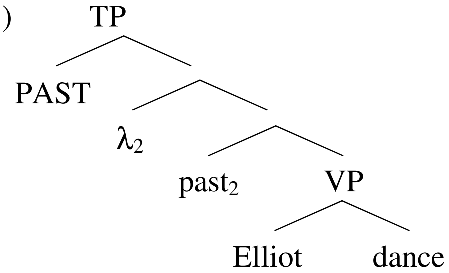
(3) $[\![\textbf{dance}]\!]^g = \lambda x\in D_e.\ [\lambda t\in D_i.\ [\lambda w\in D_s.\ \text{x danced at t in w}]]$

The past tense morphology (represented as $past_2$ in (5) below) translates to a variable of type i, whose value is given by a variable assignment. Syntactically, it is generated right above VPs, perhaps as the head of T. This past morpheme itself does not have meaning of anteriority. It is given by the phonetically null operator PAST, whose denotation is as follows:

(4) $[\![\textbf{PAST}]\!]^g = \lambda P\in D_{<i,<s,t.>>}.\ [\lambda t\in D_i.\ [\lambda w\in D_s.\ \text{there is a time t' such that t'} < \text{t}$ $\text{and that P(t')(w) = 1}]]$

Now, in order to have a proper interpretation, we need to insert a λ abstractor above past$_2$, and the whole structure looks like this:

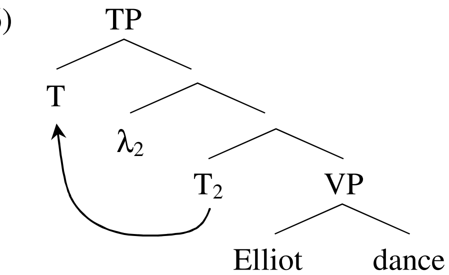
(5)



While Kusumoto's model works nicely, the formation of the configuration in (5) looks like a stipulation. Why is there a phonetically null operator? And how does the λ get inserted? If past$_2$ is the past tense morpheme, how come it is outside the VP and yet able to get affixed onto *dance*? Now, remember the discussion in the last part of section 1.2. In Heim & Kratzer's system, variable binding is achieved by means of movement. Since PAST is a quantifier of type <i,<s,t>>, and past$_2$ is interpreted as a variable of type i, once we assume that past$_2$ is the trace of the movement of PAST, the configuration in (5) is quite straightforwardly accounted for.

In Kusumoto's theory, PAST is a phonetically null operator that might be the T head and past$_2$ is the past tense morpheme. Now if we suppose that PAST is actually the T head, then it means T has originated in the position where the past$_2$ is:

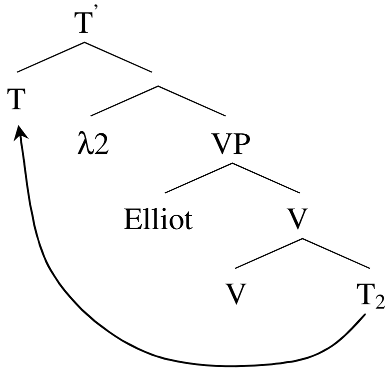
(6)



However, this is quite a strange movement. It is quite mysterious why T has to move in this construction. First, there is no type mismatch prior to the movement, and second, the landing site of the movement is right above the extraction site, resulting in no semantic effects. Such movement is redundant, must be excluded for the sake of economical computation. In addition, the problem the V and T (tense morpheme) are not adjacent still remains.

To solve these problems, I propose that T originates as the sister of T and moves out because of a semantic type mismatch (this possibility is also suggested in von Fintel's (2001) class notes).

(7)



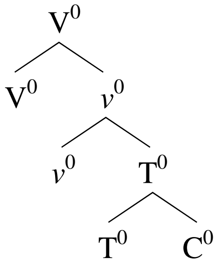
This way, V and T are next to each other to begin with, and so we can naturally account for the morphological merger of V and T. Now it is also clear why PAST is phonetically empty in Kusumoto's analysis. PAST and $past_2$ are actually two copies of the same thing, namely T, whose phonetic content is realized as the past tense morpheme. Since one morpheme should not be pronounced in different places, one of them, i.e., the PAST position in this case has to be phonetically null. This is not always the case: In French, the PAST position is pronounced, by sort of pied-piping the phonetic content of the V that T has been affixed to, yielding so-called V to T raising. Note that here we do not have movement to a non-c-commanding position, a problem that has been present in previous analyses of head movement.

2.2 Formation of a complex head

The idea that T originates as the sister of V and that is what is behind V to T raising or T to V lowering leads to a drastic reformulation of syntactic theory. If the T head originates as the sister of the V head, such configuration must be generalized to all heads, since there should be no reason that only T are V have to be treated specially in syntax. In addition, head movement is not limited to between V and T. V to C movement is seen in interrogative sentences in English and French, and is also often exhibited in V2 languages. So it is natural to think that C originates as the sister of T. In this vein, I propose that heads constitute a complex head by binary merge. Also, there is an independent support for this view, since this is the only way to avoid the problem that previous analyses of head movement and Matushansky's theory share. Recall that models where a moved head and a target head are merged after the head movement require access to the target head, which is not the root of the tree. This unwanted conclusion is inevitable as long as we assume that the merger takes place after the movement. The only logically possibility that gets around this, then, is to assume that heads are merged before they undergo head movement.

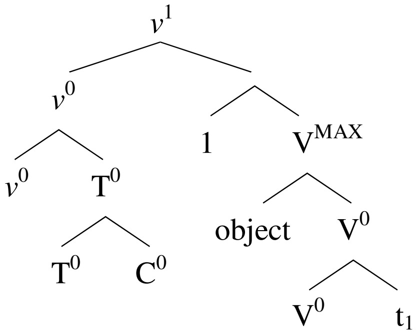
To illustrate my idea, let us think of a simple sentence whose only heads are C, T, *v* and V, disregarding the subject and object noun phrases. The derivation of this sentence begins with forming the head complex which itself is a head in (8).

(8)



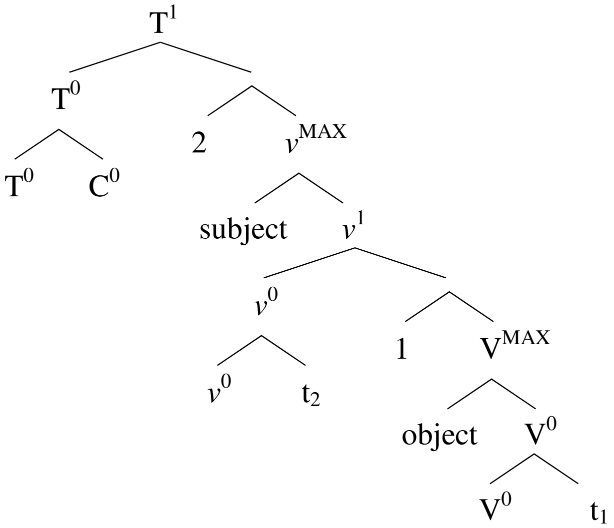
Next, the object (if any) is merged and V gets projected to form a projection of V. Then, the head complex dominated by v^0 moves by head movement, and v^1 is formed.

(9)



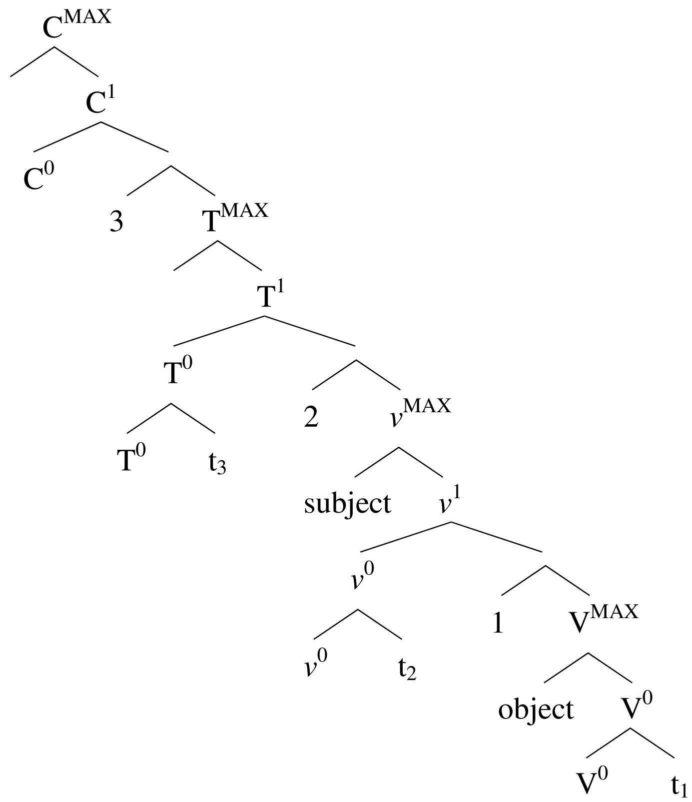
Next, the subject is merged and the head complex dominated by T^0 moves by head movement.

(10)



If necessary, the subject moves to SpecTP by A movement (here I am not showing that for simplicity's sake), and then C^0 moves by head movement.

(11)



θ-role assignment and Case assignment as standardly assumed takes place at the relevant steps in the course of the derivation.

If T^0 is a quantifier and it has to move to avoid a semantic type mismatch, it is natural to think that all other moving heads move for the same reason. That is, they are all quantifiers and move because of a type mismatch. I propose that v^0 is quantifier over events, and C^0 is quantifier over possible worlds. I develop and discuss their semantics in Section 3 extensively.

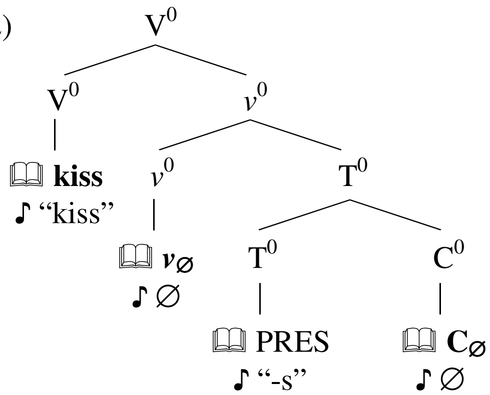
In this derivation, what is considered to be the complement to a head except for V is a node from inside which the head has moved by head movement. In X-bar theory, the specifier and the complement are defined as structurally different positions occupying in the X-bar schema. However, in the bare phrase structure proposed in the minimalist program, no such structural difference is maintained. In bare phrase structure, both the complement and the specifier are non-adjunct that are not projected. The only difference between the two is that the complement is the first of such nodes and the specifier is merged after the complement. Thus, in the minimalist framework, the difference between the complement and the specifier appears to be rather arbitrary. Now, if we define the complement as a node from within which its head-to-be is moved by head movement, the arbitrary distinction between the complement and the specifier will disappear. That is, all externally merged nodes will be specifiers, regardless of whether they are the first to be merged with a head. On the other hand, complements are formed only as a result of head movement. If we take this line, the object in the above derivation will be a specifier and not a complement of the V^0. Since specifiers are assumed to be merged to the left, in order to achieve the Verb-Object order in English, one need to assume that the phonetic content of V^0 is dragged to the v^0 position, as I explain in the next section.

2.3. Deriving various placements of heads

In this section, I show how we may derive head movement in the traditional sense (i.e., phonetic displacements of heads). Chomsky (1994, 1995) proposes that Morphology deals only with X^0, and that it gives no output if presented with an element that is not X^0. An essential idea behind this is that only heads may be morphologically merged onto another lexical item. Given this, in my model, morphological merger is possible within a head complex since all the nodes in it are heads. Thus, if a head complex involves some bound morphemes, they will phonetically be fuses with adjacent heads in the same head complex.

Now, let us take sentences (1b) and (1c) from Section 1.1. Since the derivation starts by forming a head complex, we first create the structure in (12) for *John often kisses Mary*.

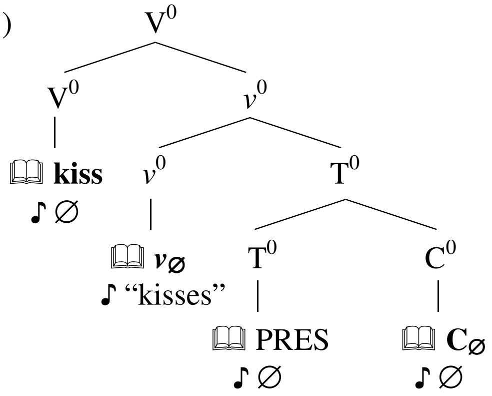
(12)



Here, 🕮 indicates the meaning (i.e., lexical entry) of a head, and ♪ represents its sound. 🕮 **X**$_\varnothing$ represents the semantics of a phonetically null X head, and PRES is short for PRESENT. We know that V^0 and T^0 must be morphologically merged, and so v^0, being sandwiched by V^0 and T^0, must also join them (i.e., v^0 is a bound morpheme that must be phonetically fused with the verb). Now, let's assume such merger takes place before any head movement occurs. Assuming that the same phonetic content cannot be pronounced in multiple positions, we need to determine where the phonetic content of the resultant inflected verb should be.

Now, if we assume that the English v^0 has the property of attracting the merged phonetic content to its position, we will obtain the following:

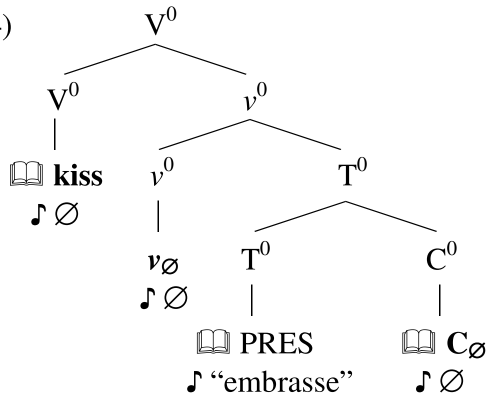
(13)



Note that here, only the phonetic content moves, but the meaning of each head remains in its "home" position. If the derivation proceeds as I illustrated in the previous section, we will obtain the English order, with the seeming effect of T lowering onto V.

In the case of French, instead of v^0, T^0 has the property of attracting the merged phonetic content. Hence:

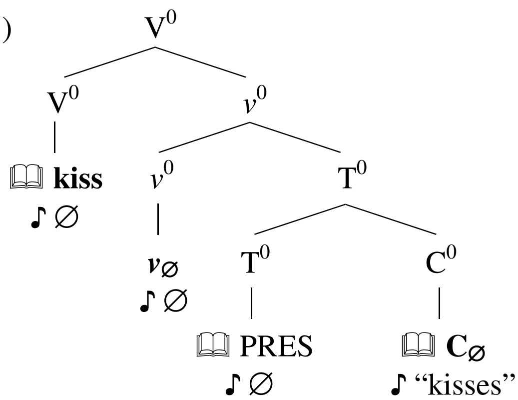
(14)



When the derivation is over, we will see *embrasse* in the T position.

For V2 languages, we can think that in addition to V^0, v^0 and T^0, phonetically null C^0 also needs to be merged together, and that C^0 has the property of attracting the merged phonetic content. Then, we obtain the following (pretending that English is V2):

(15)



After head movement, then, we will obtain the V2 order.

In this account, the meaning of each head never leaves its "home" position regardless of whether its phonetic content moves. Therefore, the English and French sentences have exactly the same meaning, although the verb appears in different positions. In this respect, head movement is similar to such A-bar movement as *wh* movement and QR. Although *Wh* movement and QR are supposed to take place in all languages in order to obtain the proper interpretation of the sentence, such movement is not always overtly seen as phonetic movement. The apparent raising of a head could then be considered as pied-piping, a phenomenon often seen with *wh* movement.

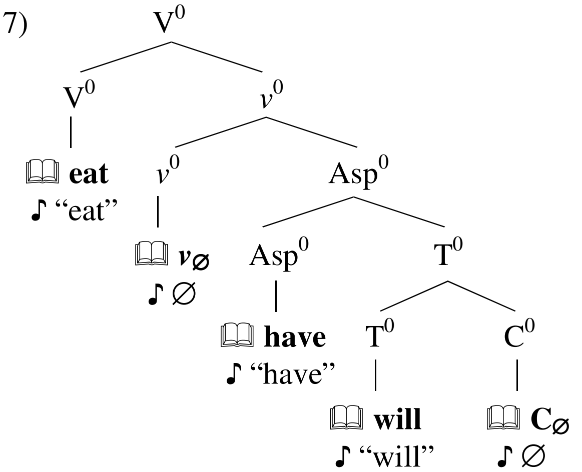
It should be noted that Travis's Head Movement Constraint should entail an adjacency condition of heads for morphological merger in my model. Consider the sentences in (16). Although (16a) is well-formed, the derived (16b) is ungrammatical, since it would violate Head Movement Constrain by moving the head *have* to C past the position occupied by the head *will* by head movement.

(16) a. Mary will have eaten the cherries.

b. Have Mary will t eaten the cherries?

In my model, (16b) could be generated without actually involving long distance head movement. Let's assume that *will* is generated as a T head, and *have* is generated as an Aspect head as follows:

(17)



Then, (16b) would be generated if *have* were able to be morphologically merged onto C^0, without affecting the *will* in between. Therefore, morphological merger of heads should be possible only when they are adjacent so that Head Movement Constraint may not be violated.

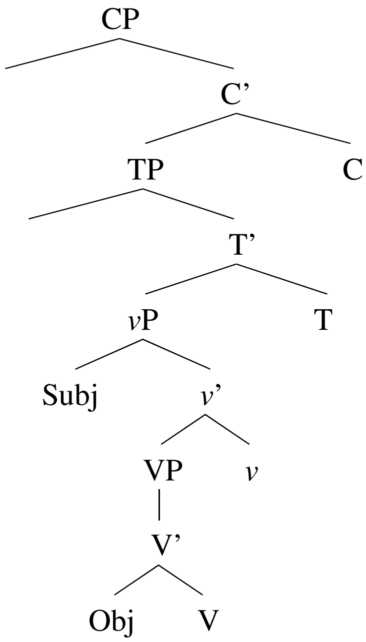
Lastly, I point out a problem with my model of the derivation. In my model, if some head other than C attracts the phonetic content of C, it is expected that the complementizer would phonetically appear in the middle of a clause, rather than at the edge of a clause, but we do not seem to find such cases. However, given that the phonetic movement takes place only when morphological merger happens, even if such movement of C takes place, the chances are that the complementizer is not recognized as a complementizer. David Pesetsky (p.c.) suggests that the *-ing* morphology in English might be a C head that has undergone such phonetic movement in some cases. Similarly, the infinitival *to* in English might be analyzed as the fusion of a nonfinite T and a C.

2.4 Do we need head directionality?

2.4.1 Straight forward "head-final languages" like Japanese and Korean

In English, the complement generally follows the head. Thus, it has the SVO order and prepositions. By contrast, languages like Japanese and Korean have the SOV order and have postpositions rather then prepositions. The fact these languages have a mirror-image word order of the word order of languages like English has often been accounted for by a binary head directionality parameter. In Japanese and Korean, the complement generally precedes the head, so these languages' head directionality parameter would be set to head-final. If this is the case, the tree of a sentence whose main predicate is a verb would be something like the following (omitting possible movement of the subject and the object):

(18)



Then, whether or not there is head movement taking place, it is expected that V, *v*, T and C would be seen next to one another in the sentence (and presumably in this order). This prediction is actually borne out. Consider the following Japanese sentence:

(19) Taro-wa [$_{CP}$ Hanako-ga ringo-o tabe-ta-to] omo-u
Taro-TOP Hanako-NOM apple-ACC eat-PAST-COMP think-PRES

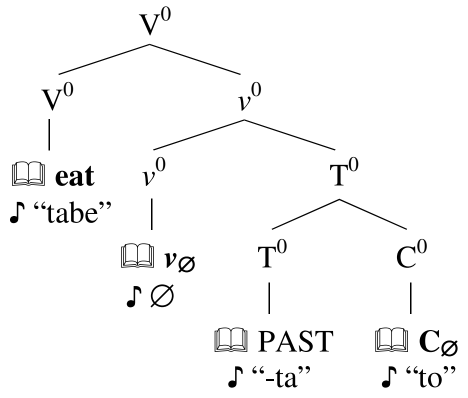
"Taro thinks that Hanako ate an apple."

Here, *ta* would be analyzed as T and *to* as C. Assuming that the *v* head in this particular case is phonetically empty, we see a V-*v*-T-C sequence in the embedded clause in (19).

Now, if the head directionality parameter existed in my framework, it would be translated to the parameter that decides whether head movement moves a head to the left or the right, since the complement of a head is a constituent out of which the head has moved by head movement. If such a parameter existed, anything regarding world order that has been accounted for by the head directionality parameter in the conventional model would have an account in my framework in the obvious translation.

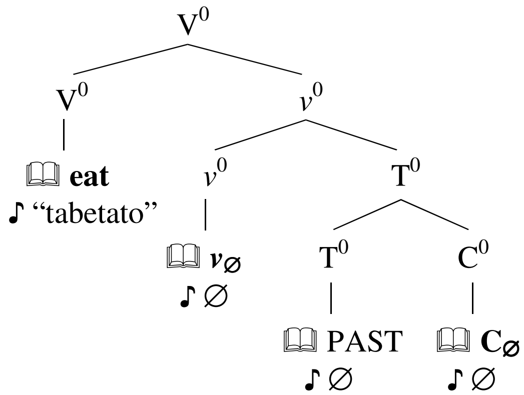
In my model, however, the word order of languages like Japanese and Korean could be derived without resorting to such a parameter. To illustrate this, let us take the derivation of (19) for example. For the embedded clause of (19), the head complex in (20) should be formed:

(20)



ta and *to* In Japanese are usually not considered as independent words in the sense that they always need another word to hang on to. Thus, we could think that they are bound morphemes and hence should be morphologically merged when the head complex is formed. If we assume that V^0 has the ability to attract the fused phonetic content, we obtain the following:

(21)



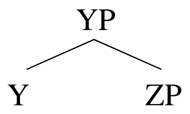
Then, after the derivation of the embedded clause is over, we correctly obtain the order in which the object *ringo-o* precedes *tabetato*.

2.4.2. Kayne's approach

Kayne (1994) proposes that word order is determined by antisymmetric c-command. On this approach, every language has the specifier-head-complement order (i.e., "head-initial"), and so the head directionality parameter cannot be maintained. Thus, if a language appears to exhibit a word order in which the complement precedes the head, it is actually a result of movement (see Koopman (2003) for arguments for the Kaynian approach to Korean and Japanese).

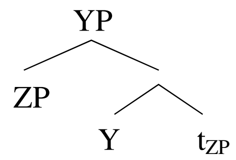
Let us see how such an order can be derived under this approach. In (a), ZP is the complement of YP, so it follows Y:

(22)



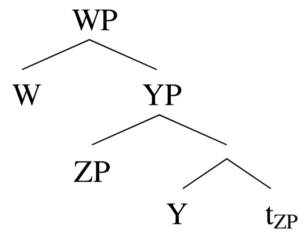
In order to achieve the complement-head order, ZP moves to a specifier of YP:

(23)



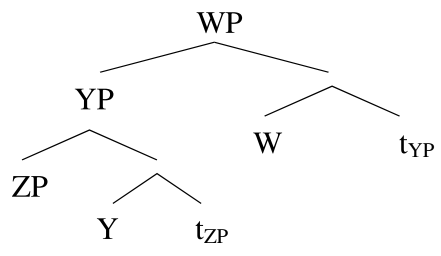
Now, YP is merged with W, and since YP becomes the complement, it follows W:

(24)



Finally, in order to achieve the same surface order as in (18), YP moves to a specifier of XP:

(25)



Thus, the complement of a given phrase always moves to a specifier position of the same phrase to achieve the surface specifier-complement-head order.

Kayne's theory and the proposal that I made in the previous subsections share the idea that every language has the specifier-head-complement and so the head directionality parameter does not exist. However, the way in which the "head-initial" and "head-final" orders are derives is quite different in the two models. In Kayne's theory, the "head-initial" order is the base word order without movement, and the "head-final" word order is derived by successive movement of complement to specifiers. By contrast, in my model, the noun phrase object is not a complement since the complement of a head is a constituent from which the head moves out by head movement. The "head-initial" word order is derived by the fact that the phonetic content of V is pied-piped by another head such as *v*. In the previous subsections, I further suggested that the "head-final" order may be achieved if V attracted the phonetic content of all the heads in the same head complex.

Unlike my proposal, Kayne's theory suffers from one obvious drawback, however. In (25), ZP intervenes between Y and W. In agglutinating languages like Japanese, Y and W together behave as if they were one word, but ZP is an independent word by itself. If ZP stays in WP and if Y and W are to morphologically become one word, it is expected that the intervening ZP would also become part of the resulting word, but that is not the case. Moreover, as I mentioned above, only heads should be able to be merged morphologically, but in (25), neither the Y head nor the W head is directly merged with the other or other's projection. Thus, the configuration in (25) implies that Y and W will not become one word. By contrast, in my proposal, heads that end up acting as one morphological word start together as a head complex. Therefore, the morphological merger may quite naturally take place.

Let me lastly note that Kayne's approach is not itself essentially incompatible with my framework, although if we adopted Kayne's approach in conjunction with my framework to derive the "head-final" order, the derivation would be necessarily extremely complex. Since my approach has the same specifier-head-complement order, we can in principle move complements to specifiers to achieve the "head-final" word order. In my framework, although complements do not include noun phrase objects, they should still be able to undergo movement to a specifier position, since we see movement such as VP topicalization in any case.

In the following subsections, I present more data and analyses that are compatible with my proposal.

2.4.3 Mixed directionality–German and Hindi

In this subsection, we look at languages that appear to exhibit mixed head directionality, namely German and Hindi. Usually, German and Hindi both exhibit the Object-Verb order, which appears to suggest that they are head-final languages. Strangely enough, though, in these languages, the finite clausal object of a verb follows the verb (except when the clausal complement is moved by topicalization, of course), and the complementizer of a finite clause appears in the beginning of the clause and not at the end:

(26)	a.	raam-ne	sochaa	ki	mohan	hoshiyaar	hE.	(Hindi)
		Raam-erg	thought	that	Mohan	smart	is	

"Raam thought that Mohan is smart."

b. Hans sagte daß er keine Zeit hätte. (German)
Hans said that he no time had

"Hans said that he had not time."

(Bayer (1996), p. 191)

Thus, the complementizer and the verb taking the finite clausal complement have their complement to their right, but these characteristics are typical of head-initial languages. In addition, while Hindi has postpositions, German has prepositions. Thus, Hindi and German do not seem to have a fixed head-directionality, and if these orders all reflect the base-generated orders, the Principles and Parameters approach with a binary head-directionality parameter will be untenable.

What makes the situation even more complex in these languages is the fact that if there is an auxiliary verb following a verb that takes a finite clausal complement, the clausal complement appears to the right of not only the verb but also the auxiliary verb as seen below:

(27) a. raam-ne socaa thaa ki mohan hosiyaar hE. (Hindi)
Ramm-erg thought be(pst) that Mohan smart is

"Raam thought that Mohan is smart"

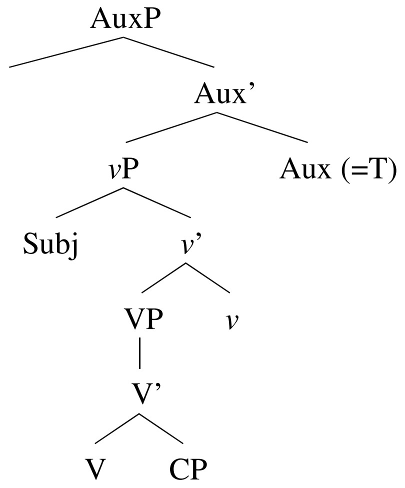
(Mahajan (1990), p. 127)

b. dass Stefan gesagt hat, dass er Äpfel gegessen hat. (German)
that Stefan said has that he apples eaten has.

"that Stefan said he ate apples."

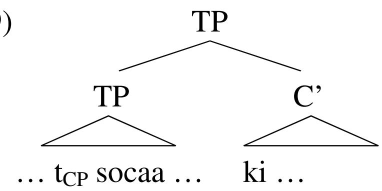
The fact that the auxiliary verb comes after the verb is expected if we assume a head-final tree such as (18) above. However, the fact that the clausal complement follows the auxiliary verb is not. Since the auxiliary verb should be analyzed as the T head or else some other functional head above *v*P, even if we assumed that the matrix verb takes its clausal complement to its right, the clausal complement would still precede the auxiliary verb as illustrated by the following tree.

(28)



Mahajan (1990), who assumes that Hindi is uniformly head-final, analyzes that the finite embedded clause in Hindi is generated to the left of the verb just like non-clausal objects and moves rightward and adjoins to matrix IP (= TP in (24)).

(29)



Here and below, I replace Mahajan's I's with T's, assuming that the difference between I and T is only nomenclatural, since I am using T instead of I in this paper. Thus, the embedded clause moves to the right across the auxiliary *thaa* to achieve the Verb-Aux-Embedded Clause order. Similarly, one might analyze the finite embedded clausal complement in German as being adjoined to TP.

This extraposition analysis is problematic, however, since the indirect object seems to c-command the embedded clause, as discussed for German by Bayer (1996). (30), taken from Bayer (1996) (p. 196), shows a bound pronoun reading and a Condition C violation:

(30) a. daß der Direktor [jeder Putzfrau]$_i$ persönlich mitteilte
that the director each cleaning-lady personally told

[daß sie$_i$ entlassen sei]
thet she fired was

"that the director personally told each cleaning-lady that she was fired."

b. *daß der Direktor ihr$_i$ persönlich mitteilte [daß die Putzfrau$_i$ entlassen sei]
her the cleaning-lady

(31) shows a bound pronoun reading and a Condition C violation for similar Hindi constructions:

(31) a. raam$_i$ [apnii$_i$ har behen ko]$_j$ kehtaa hE ki woh$_j$ achchii hE.
self's every sister say she good

"Raam says to every sister of his that she is good."

b. maayaa usko$_{i/*j}$ kehtii hE ki raam$_j$ achchaa hE.
Maayaa to him say is that Raam good is

"Maayaa says to him that Raam is good."

These data show that the indirect object somehow c-commands the embedded clause.

An obvious hypothesis is that the extraposed embedded clause is reconstructed at LF, so that the indirect object c-commands the pronoun in the embedded clause at LF. This seems unlikely given the sentences like the following:

(32) a. daß der Direktor es [jeder Putzfrau]$_i$ persönlich mitteilte [daß sie$_i$ entlassen sei]

b. *daß der Direktor {es ihr$_i$ / ihr$_i$ es} persönlich mitteilte [daß die Putzfrau $_i$ entlassen sei]

(33) a. Raam$_i$ [apnii$_i$ har behen ko]$_j$ yeh kehtaa hai ki woh$_j$ achchii hai.

b. Maya usko$_{i/*j}$ yeh kehtii hai ki Raam$_j$ achchaa hai.

These are the same as the sentences in (30) and (31) except that the canonical object position of the matrix clause is filled with the expletive *es* for German and *yeh* for Hindi respectively. Still, these have exactly the same judgments as their corresponding sentences above. When the canonical object position is already filled by something else, it is hard to believe that the embedded clause is generated there in (32) and (33). Mahajan (1990) proposes that when the expletive object *yeh* is present, the embedded clause is generated in an IP-adjoining position, co-indexed with the expletive. If we take this line of approach, there could be by definition no such thing as reconstruction of the embedded clause. Then, since we need to be able to account for (32) and (33) anyway, we should think of an account without resorting to reconstruction, and if once we have an account for (32) and (33), that would take care of (30) and (31) as well.

In any case, the embedded clause in the above sentences is in an adjoining position to TP. Therefore, in order for the indirect object to c-command the embedded clause, the indirect object has to be raised to a position higher then the embedded clause. However, this is also problematic.

First, if DPs in this embedded clause can enter crossover configurations, the raising of the indirect object will cause a weak crossover, since the indirect object that is co-indexed with the pronoun in the embedded clause moves from a position that does not c-command the pronoun to a position that does.

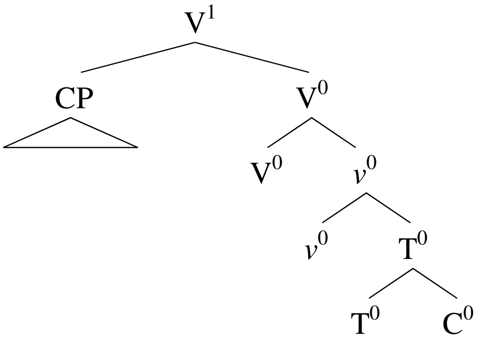
Even if we assume that the weak crossover does not obtain, the analysis is still problematic. In (31a) and (32b), *apnii* is a reflexive possessive pronoun co-indexed with the subject and therefore must be A-bound by the subject *Parmesh* or its trace. Thus, both the subject and the indirect object have to be in A positions. Now, if the embedded clause is outside the innermost TP as in (29), both the subject and the indirect object should have moved outside the innermost TP. But then, it is not clear at all why there are A positions available in that region. If the embedded clause is actually adjoined to T' and so inside TP, the A position for the subject is guaranteed, but it is not clear why an A position for the indirect object exists below the subject A position and above the embedded clause. In either case, we have no explanation of why the indirect object (and the subject, depending on the analysis) should move.

Thus, German and Hindi pose many difficult problems regarding word order. The Kaynian approach does not really help. Since the embedded clause follows the auxiliary verb, we would have the same conclusion that the embedded clause is higher then the T head (Suppose that Y = V (or *v*) and W = T (Aux) in (25) above. Then, the embedded clause would be adjoined to WP).

2.4.4 Embedded clauses

Here, I propose a solution to the problem on embedded clauses in German and Hindi. First, let us suppose that that the embedded CP is computed first and then it is merged with the matrix V. This might seem like an ordinary assumption, but recall that in my model, the embedded CP would then not be the complement of the matrix V, since the complement of V should be the constituent out of which V has moved by head movement. In other words, the embedded CP would then be a specifier:

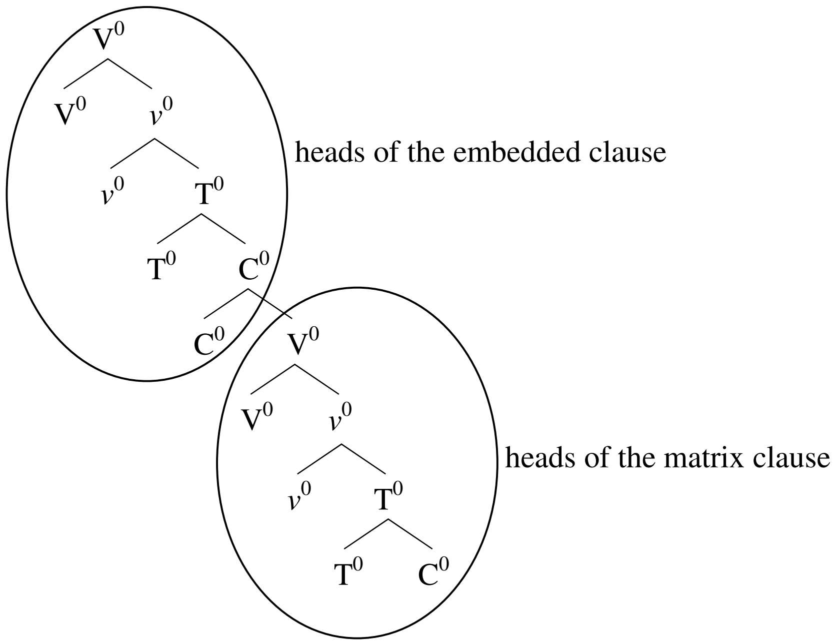
(34)



However, since Huang (1982), it has been known that movement out of a specifier is in general impossible. Therefore, extraction out of the embedded CP should be impossible, but this is clearly an unwanted conclusion.

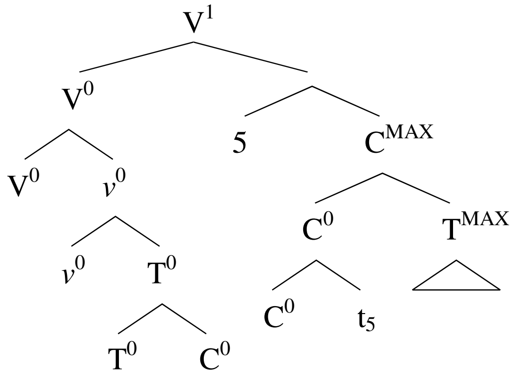
Thus, in order for the embedded CP to be the complement, it should be the case that the matrix V has moved from inside the embedded clause by head movement. More precisely, the moved V head is generated as the sister of the embedded terminal C head, and this means that this V head is part of the head complex for the embedded clause. Since this V head is itself a complex head that consists of the heads of the matrix clause, it is concluded that the head complex of the matrix clause and that of the embedded clause should be formed together as a big head complex in the beginning of the derivation as illustrated below:

(35)



This head complex is formed by merging the heads of the embedded clause after the head complex for the matrix clause is formed. After the derivation of the embedded clause is over and the matrix V head has moved, the tree would look like the following:

(36)



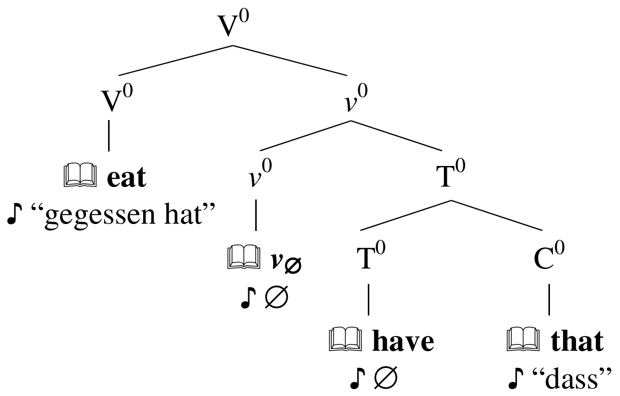
The embedded clause is now the complement of the matrix V, so extraction out of it is possible.

Semantically speaking, this derivation makes sense. Schlenker (200?) argues that propositional attitude verbs quantify over contexts. Following Schlenker, Heim (200?) also proposes a syntax in which an attitude verb is generated inside its clausal complement and moves out to remedy a type mismatch. I discuss the semantics of my model in more detail in Section 3.

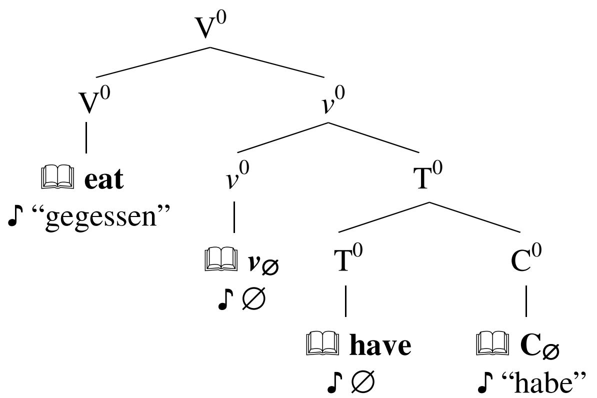
I now propose that this derivation can actually account for the word order in German and Hindi we looked at in the previous subsection.

Let us first consider German. German is a V2 language, and this means that in my model, when the head complex is formed, it is required that the C head position be phonetically not empty, as I briefly discussed in Section 2.3. Thus, if the C head is a phonetically non-empty C such as *dass*, the C head position does not attract the phonetic content of any other heads. By contrast, if the C head position is phonetically null, then it attracts the nearest pronounceable chunk, i.e., either the auxiliary verb or the inflected verb. On the other hand, in order to achieve the "head-final" order of the V-T sequence, the phonetic content of whatever is left is attracted to the V head.

(37)a. When the C head is phonetically non-empty

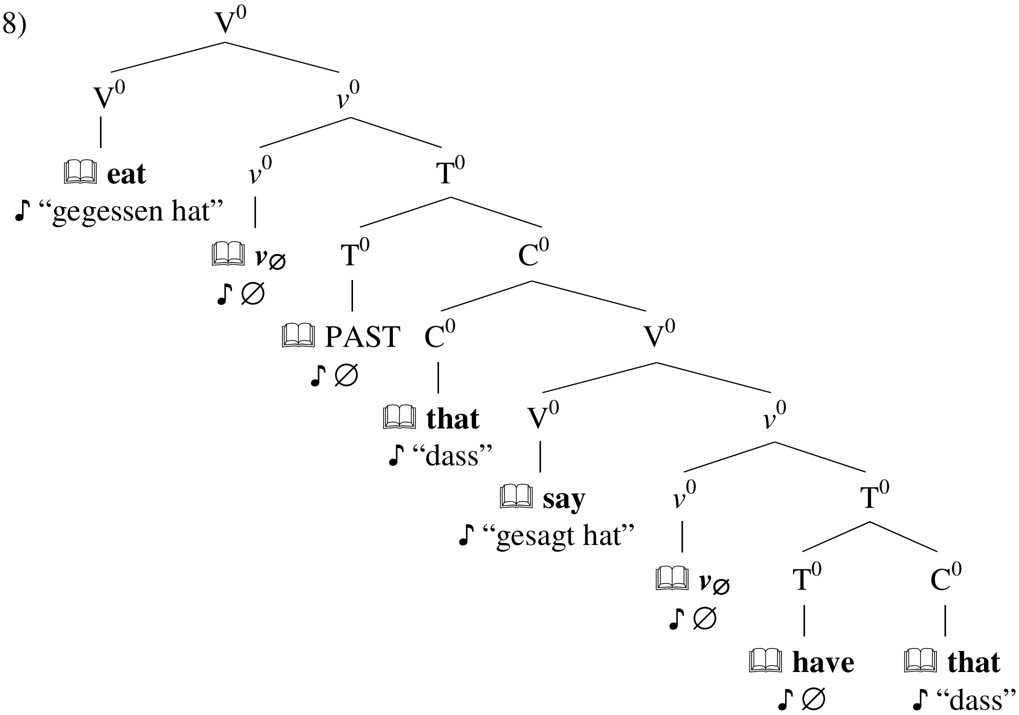


b. When the C head is phonetically empty



When the derivation is complete, these lead to structures with the desired word orders. This explains the problematic placement of the embedded clause seen in the previous subsection as well. When a verb takes a clausal complement, a head complex like (35) is formed. After the phonetic content has moved so that the C will be phonetically non-empty and that the rest gathers at V in each clause, we obtain the following:

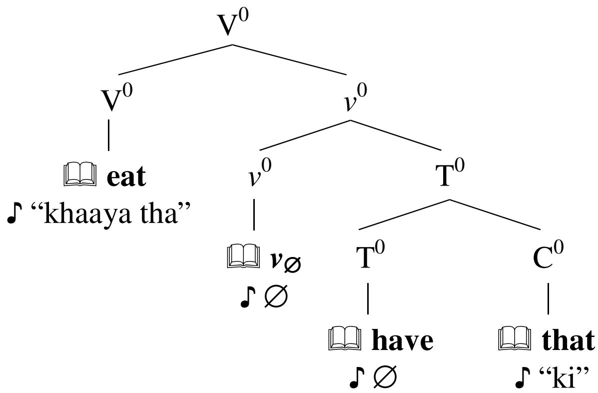
(38)



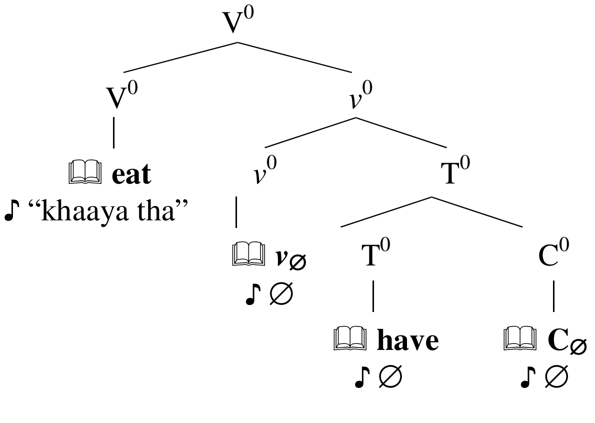
Again, when the derivation is complete, we obtain the desired word order.

Unlike German, Hindi is not a V2 language. So whether or not the C head is phonetically non-empty, it does not attract the phonetic content of the neighboring heads in the head complex. Like German, however, Hindi's V attracts the phonetic content of the heads in the same clause except for C, since C is head-initial in Hindi, i.e., C appears in the beginning of the clause:

(39)a. When the C head is phonetically non-empty



b. When the C head is phonetically empty

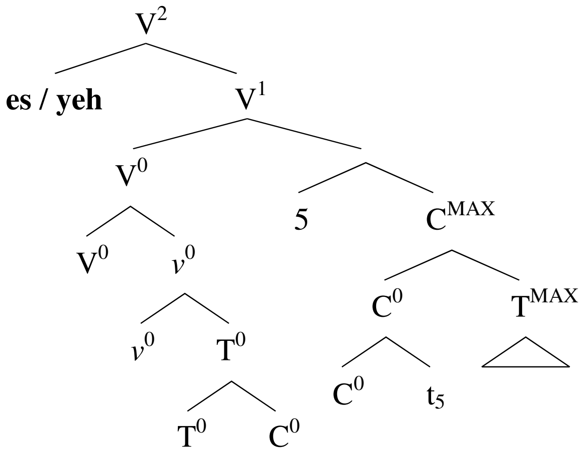


Given the above characteristics, the placement of the embedded clause can be derived pretty much the same way as the German example above.

Finally, on the approach that does not resort to a head-directionality parameter, the fact that German has prepositions and Hind has postpositions will be accounted for by morphological attracting property of some heads involved in the structure.

The analysis that I have presented above naturally obviates the c-command problem. The embedded CP actually ends up being the canonical complement position of the verb as in (36), and therefore the indirect object would naturally c-command the embedded clause. Moreover, the expletive object *es* / *yeh* can just be in the “canonical object position” in my framework, namely the specifier of V:

(40)



Thus, although it is not clear why the expletive object is optional, its existence would not interfere with the embedded CP, the complement of V. In the rest of the paper, I take the analysis I have developed here to be correct and assume that a propositional attitude verb with a clausal complement is derived from inside the clausal complement.

2.4.5 Embedded clauses in consistently “head-final” languages

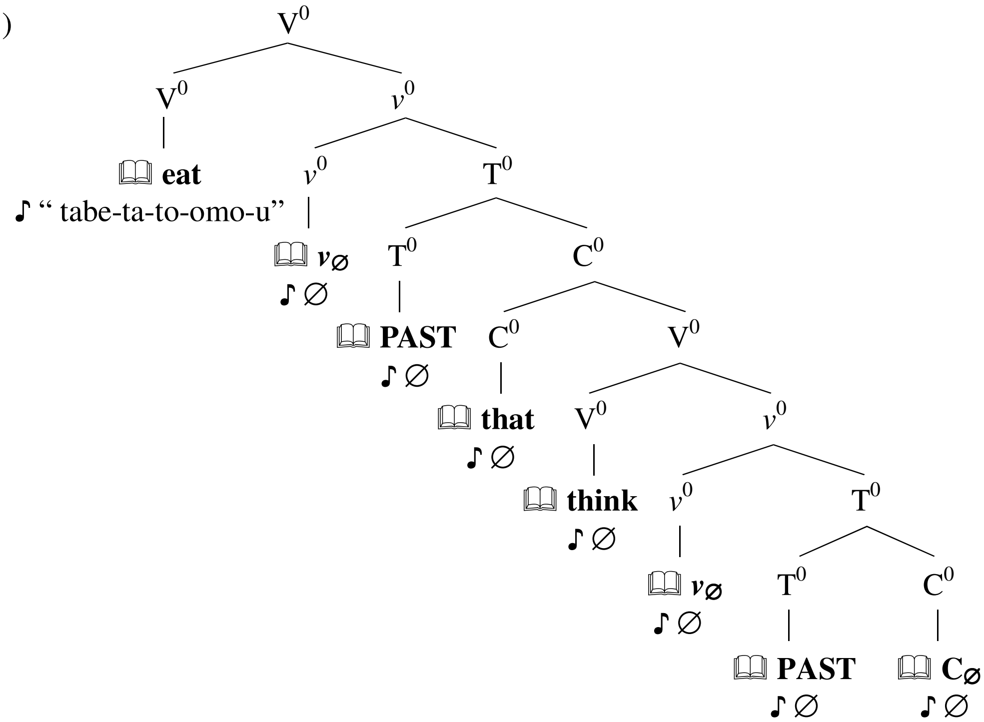
In consistently “head-final” languages like Japanese, a clausal complement precedes the verb just like a noun object does. Below, I repeat (19) above:

(19) Taro-wa [$_{CP}$ Hanako-ga ringo-o tabe-ta-to] omo-u
Taro-TOP Hanako-NOM apple-ACC eat-PAST-COMP think-PRES

“Taro thinks that Hanako ate an apple.”

If the heads of an embedded clausal complement are part of the head complex of the higher clause as I proposed in the previous section, how can we derive this word order of consistently “head-final” languages? One hypothesis is that the embedded V attracts the phonetic context of all the heads including the heads of the matrix clause as illustrated as follows:

(41)



However, it appears sentences like the following are difficult to account for with this hypothesis:

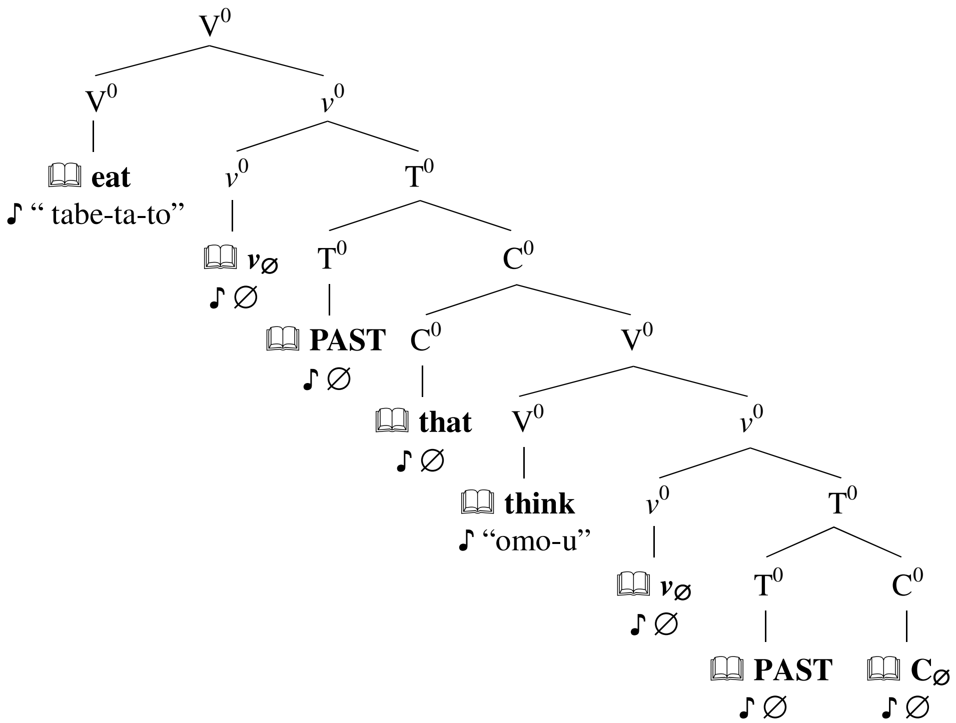
(42) Taro-wa [$_{CP}$ Hanako-ga ringo-o tabe-ta-to] tsuyoku omo-u
Taro-TOP Hanako-NOM apple-ACC eat-PAST-COMP strongly think-PRES

“Taro strongly thinks that Hanako ate an apple.”

Here, the adverb *tsuyoku* “strongly” appears between the embedded predicate and the matrix predicate. If the embedded V attracts the phonetic content of all heads as in (41), this cannot be accounted for.

An alternative analysis is to assume that the derivation in consistently "head-final languages is similar to the derivation in German and Hindi in that the embedded V and the matrix V attracts the phonetic content of their respective clauses as follows:

(43)



Given that we do not seem to see cases where a head that belongs to an embedded clause appears in the matrix clause or a head that belongs to the matrix clause appears in an embedded clause, there seems to be a condition that the phonetic movement of heads be "bounded". That is, even though the heads of the matrix clause and those of an embedded clausal complement may form a big head complex in the beginning of the derivation, a head is allowed to move phonetically only to heads that belong to the same clause as it does. If so, (43) will be more plausible than (41).

When the head movement of the matrix V takes place, the matrix verb will appear to precede the embedded clause. Thus, the embedded CP would have undergo movement to the left of the matrix verb. Postulating movement of the embedded CP is not problematic by itself. Since the adverb intervenes the object and the verb in (42), as long as we assume that the object is generated right next to the verb, such movement might be necessary in any case.

Since I have not reached any decisive conclusions regarding the word order of consistently "head-final" languages, I leave this issue open for future research.

Chapter 3 Semantic Composition and Presuppositionality

In this chapter, I develop the semantics of my framework, where moving heads denote quantifiers. I also investigate the internal structure of noun phrases and propose different syntactic structures for presuppositional and cardinal noun phrases. More precisely, I argue that presuppositional noun phrases have quantifiers inside them, whereas cardinal noun phrases have an anaphor that denotes a situation.

3.1 C as a quantifier over possible worlds

I begin this chapter with a discussion on the semantics of the complementizer. As I briefly mentioned in the previous chapter, I propose that C denotes a quantifier over possible worlds, and following Schlenker (200?) I also propose that attitude verbs denote quantifiers over contexts and define a relevant Intensional Functional Application rule.

The semantics of matrix clauses is usually evaluated with respect to the "actual world", and this "actual world" has often been assumed to be a single possible world. Stalnakers (1984) argues that the discourse of a conversation can be represented as a context set, the set of possible worlds that represents that context. More precisely, in the possible worlds in the context set, all and only the propositions that have been established to be true in the discourse are true. Every time a participant in the conversation learns some fact, he or she gets rid of all possible worlds in the context set in which the learned fact does not obtain. A conversation is then a process of reducing the context set.

Given this idea, unless we are omniscient, what we conceive of as the "actual world" may not actually be a single world, but the set of all and only the possible worlds in which what one considers as facts hold. Then, the "actual world" with respect to which we evaluate sentences should also be the context set, rather than a single possible world. Thus, given the context set A which represents the supposed actual facts, proposition P(w) should be true in the context if and only if the following holds:

(1) $\forall w[w \in A \rightarrow P(w)]$

Now, consider (2a). I can truthfully utter this without knowing who John's wife is. I could be certain that whoever happened to be John's wife came to the party, no matter whether it was Mary or Jane. If sentences are evaluated with respect to a single possible world, there could not be more than on possibility as to the identity of John's wife. Given a single possible world, *John's wife* might have the same denotation as *Mary* in that world for instance. However, even if John's wife was actually Mary, (2b) surely does not mean the same things as (2a), when I, who didn't know who John's wife was, utter it.

(2) a. John's wife came to the party.

b. Mary came to the party.

If we assume that sentences should be evaluated against the context set, we can capture the difference between (2a) and (2b). On this view, (2a) and (2b) will roughly be analyzed as in (3a) and (3b) below respectively:

(3) a. $\forall w[w \in A \rightarrow \text{came-to-the-party}(\iota x[x \text{ is John's wife in } w])(w)]$

b. $\forall w[w \in A \rightarrow \text{came-to-the-party}(\text{Mary})(w)]$

(3a) means that if John's wife was Mary in $w_1 \in A$, then Mary came to the party in w_1, and if John's wife was Jane in $w_2 \in A$, then Jane came to the party in w_2, and so forth. Thus, (3a) does not commit to any particular identity of John's wife. (2a) could of course mean exactly the same thing as (2b) in some situations. This is because in such cases, the context set A is such that for every $w \in A$, Mary is John's wife in w. In other words, (2a) and (2b) have the same meaning if I know that Mary is John's wife.

Where, then, does the quantification over possible worlds come from? To answer this, let us first take a look at the semantics of sentences with a propositional attitude verb such as *believe* and *know*. Sentences with an attitude verb is standardly analyzed to involve quantification over possible worlds that are compatible with what somebody believes, knows, etc., and this quantification is often assumed to come from the lexical meaning of the verb. For example, Heim & Krazter (1998) gives the following denotation to the verb *know* (p 302, with trivial modifications):

(4) $[\![\textbf{know}]\!]^{g,w} = \lambda p \in D_{<s,t>}.\ [\lambda x \in D_e.\ p(w') = 1$, for all $w' \in D_s$ that are compatible with what x knows in w]

With this, the sentence in (5a) is analyzed as in (5b) under the world assignment α:

(5) a. I know that John's wife came to the party.

b. $\forall w[w \text{ is compatible with what I know in } \alpha$
$\rightarrow \text{came-to-the-party}(\iota x[x \text{ is John's wife in } w])(w)]$

If we let A' = $\{w \mid w \text{ is compatible with what I know in } \alpha\}$, (5b) can be written as follows:

(6) $\forall w[w \in A' \rightarrow \text{came-to-the-party}(\iota x[x \text{ is John's wife in } w])(w)]$

This looks very similar to (3a), and this is not without reasons.

When I utter (2a), as long as one assumes that I am not trying to tell a lie, it is understood that I am speaking of my knowledge, and therefore, uttering (2a) would be understood to involve metalinguistic "I know" in the beginning. Then, the only difference between (2a) and (5a) is regarding whether the "I know" part in the beginning of the sentence is metalinguistic or linguistic. Now, if an omniscient person exists, the context set he has internalized will be a singleton set. If we took α to be the possible world that is the only member of the context set that an omniscient person has internalized, {w | w is compatible with what I know in α} would be a description of the context set that I have internalized, and hence A' = A. (3a) and (6) would then be identical.

Thus, the analysis in (3a), or rather (1), comes natural, once we assume that every sentence involves metalinguistic "I know" in the beginning. It is then natural to think that whatever brings the quantification over possible worlds in the semantics of a sentence with an attitude verb as seen in (5b) is also responsible for the quantification over possible worlds for the matrix clause as seen in (1). In Heim and Kratzer's (1998) system, the complementizer is semantically vacuous, and the lexical meaning of attitude verbs contains the quantification. In my framework, the complementizer (i.e., a C head) comes to its position by head movement, and head movement should occur due to a semantic type mismatch. If the complementizer were semantically vacuous, it would not be clear why such a movement is necessary. Thus, on the assumption that the moving C head is a quantifier, I propose that it is the complementizer that is responsible for the quantification over possible worlds.

Given that propositional attitude verbs do not quantify over possible worlds, what do they do, then? As I briefly mentioned in Section 2.4.4, following Schlenker (200?), I assume that propositional attitude verbs are quantifiers over contexts. By contexts, Schlenker means combinations of a set of possible worlds and an assignment for indexical expressions. Since we have separated the quantification over possible worlds as part of the complementizer's denotation, what propositional attitude verbs quantify over should be mere assignments for indexical expressions. I assume that there are special constants such as NOW, HERE, I, YOU, etc. for indexical expressions such as *now*, *here*, *I*, *you*, etc., and these are in the domain of every indicial assignment. Let c be the type of indexical assignments. Then,

(7) For any $c \in D_c$, NOW, HERE, I, YOU, etc. $\in$ Dom (c)

Sentences are evaluated under an indexical assignment and a variable assignment. For any indexical assignment c and variable assignment a,

(8) a. $[\![\textbf{now}]\!]^{c,a}$ = c(NOW)

b. $[\![\textbf{here}]\!]^{c,a}$ = c(HERE)

c. $[\![\textbf{I}]\!]^{c,a}$ = c(I)

d. $[\![\textbf{you}]\!]^{c,a}$ = c(YOU)

etc.

We have to make sure that when an attitude verb is present, the indexical assignment for the clausal complement of the attitude verb may be replaced. Thus, we define Intensional Functional Application rule (IFA) as follows:

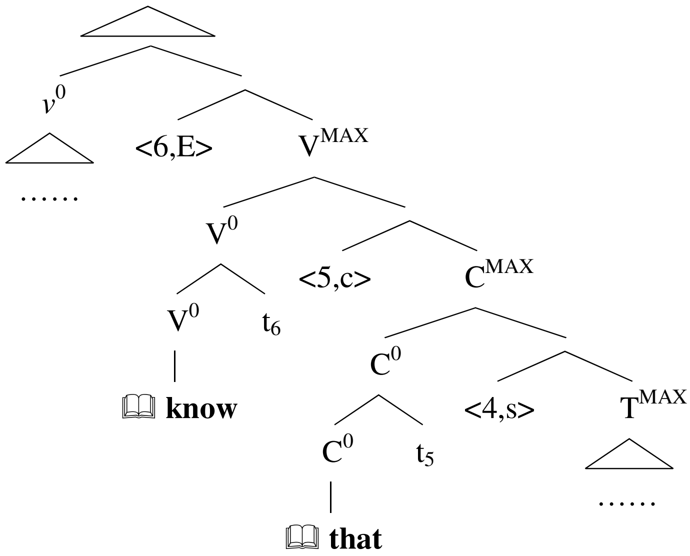
(9) Intensional Functional Application

If α is a branching node and $\{\beta, \gamma\}$ the set of its daughters, then, for any indexical assignment c and variable assignment a, if $\llbracket \beta \rrbracket^{c,a}$ is a function whose domain contains $\lambda d \in D_c. \llbracket \gamma \rrbracket^{d,a}$, then $\llbracket \alpha \rrbracket^{c,a} = \llbracket \beta \rrbracket^{c,a} (\lambda d \in D_c. \llbracket \gamma \rrbracket^{d,a})$.

As can be seen, the variable assignment remains intact when IFA is applied.

Now, let us look at a relevant tree structure with an attitude verb and discuss its semantics. As I discussed in Section 2.4.4, the attitude verb moves from the sister position of the embedded terminal C head. Thus, the tree of the sentence in (5a) will look as follows:

(10)



Here, s is the type of possible worlds, and E the type of events. Since the moved C head is a quantifier over possible worlds, it binds the trace it has left behind which is of type s. Similarly, since the moved V head is a quantifier over indexical assignments, it binds the trace it has left behind which is of type c. I am further assuming that the moved *v* head is a quantifier over events and therefore it binds the trace it has left behind which is of type E (I give the semantics of the *v* head in Section 3.3?). Given this tree structure, I propose the following lexical entries for *that* and *know*:

(11) Lexical entries

a. $\llbracket \textbf{that} \rrbracket^{c,a} = \lambda d \in D_c. [\lambda f \in D_{<c,st>}. [\lambda u \in D_{st}. \forall w \in D_s [u(w) \rightarrow f(d)(w)]]]$

b. $\llbracket \textbf{know} \rrbracket^{c,a} = \lambda e \in D_E. [\lambda \gamma \in D_{<c,<st,t>>}. \text{KNOW}(e) \wedge$
$\forall d \in D_c[\text{d is compatible with what is known in e}$
$\rightarrow \gamma(d)(\lambda w \in D_s. \text{ w is compatible with what is known in e})]]$

Let us see that the semantics of the sentence is correctly computed.

$$[\![\mathbf{<6,E>\ V^{MAX}}]\!]^{c,a}$$

$$= \lambda e\in D_E. [\![\mathbf{V^{MAX}}]\!]^{c,a^{e/<6,E>}}$$

$$= \lambda e\in D_E. [\![\mathbf{know}]\!]^{c,a^{e/<6,E>}}([\![\mathbf{t_{<6,E>}}]\!]^{g^{e/<6,E>}})([\![\mathbf{<5,c>\ C^{MAX}}]\!]^{g^{e/<6,E>}})$$

$$= \lambda e\in D_E. [\lambda e'\in D_E. [\lambda\gamma\in D_{<c,<st,t>>}. \text{KNOW}(e') \wedge \forall d\in D_c[\text{d is compatible with what is known in e'} \rightarrow \gamma(d)(\lambda w\in D_s. \text{ w is compatible with what is known in e'})]]] (e)([\![\mathbf{<5,c>\ C^{MAX}}]\!]^{c,a^{e/<6,E>}})$$

$$= \lambda e\in D_E. \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e} \rightarrow [\![\mathbf{<5,c>\ C^{MAX}}]\!]^{c,a^{e/<6,E>}}(d)(\lambda w\in D_s. \text{ w is compatible with what is known in e})]$$

$$= \lambda e\in D_E. \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e} \rightarrow [\lambda g'\in D_c. [\![\mathbf{C^{MAX}}]\!]^{c,a^{e/<6,E>,\ g'/<5,c>}}](d) (\lambda w\in D_s. \text{ w is compatible with what is known in e})]$$

$$= \lambda e\in D_E. \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e} \rightarrow [\![\mathbf{C^{MAX}}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}}(\lambda w\in D_s. \text{ w is compatible with what is known in e})]$$

Here,

$$[\![\mathbf{C^{MAX}}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}}$$

$$= [\![\mathbf{that\ t_{<5,c>}\ <4,s>\ T^{MAX}}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}} \quad \text{(by IFA)}$$

$$= [\![\mathbf{that\ t_{<5,c>}}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}}(\lambda g\in D_c. [\![\mathbf{<4,s>\ T^{MAX}}]\!]^{g,a^{e/<6,E>,\ d/<5,c>}})$$

$$= [\![\mathbf{that}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}}([\![\mathbf{t_{<5,c>}}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}})(\lambda g\in D_c. [\![\mathbf{<4,s>\ T^{MAX}}]\!]^{g,a^{e/<6,E>,\ d/<5,c>}})$$

$$= [\lambda d'\in D_c. [\lambda f\in D_{<c,st>}. [\lambda u\in D_{st}. \forall w\in D_s\ [u(w) \rightarrow f(d')(w)]]]] (d)(\lambda g\in D_c. [\![\mathbf{<4,s>\ T^{MAX}}]\!]^{g,a^{e/<6,E>,\ d/<5,c>}})$$

$$= \lambda u\in D_{st}. \forall w\in D_s\ [u(w) \rightarrow [\lambda g\in D_c. [\![\mathbf{<4,s>\ T^{MAX}}]\!]^{g,a^{e/<6,E>,\ d/<5,c>}})](d)(w)]$$

$$= \lambda u\in D_{st}. \forall w\in D_s\ [u(w) \rightarrow [\![\mathbf{<4,s>\ T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>}}(w)]$$

$$= \lambda u\in D_{st}. \forall w\in D_s\ [u(w) \rightarrow [\lambda w'\in D_s. [\![\mathbf{T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>,\ w'/<4,s>}}](w)]$$

$$= \lambda u\in D_{st}. \forall w\in D_s\ [u(w) \rightarrow [\![\mathbf{T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>,\ w/<4,s>}}]$$

Thus,

$[\![\mathbf{<6,E>\ V^{MAX}}]\!]^{c,a}$

$= \lambda e\in D_E.\ \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e}$
$\quad \rightarrow [\![\mathbf{C^{MAX}}]\!]^{c,a^{e/<6,E>,\ d/<5,c>}}(\lambda w\in D_s.\ \text{w is compatible with what is known in e})]$

$= \lambda e\in D_E.\ \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e}$
$\quad \rightarrow [\lambda u\in D_{st}.\ \forall w\in D_s\ [u(w) \rightarrow [\![\mathbf{T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>,\ w/<4,s>}}]]$
$\quad\quad (\lambda w\in D_s.\ \text{w is compatible with what is known in e})]$

$= \lambda e\in D_E.\ \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e}$
$\quad \rightarrow \forall w\in D_s\ [[\lambda w'\in D_s.\ \text{w' is compatible with what is known in e}](w)$
$\quad\quad \rightarrow [\![\mathbf{T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>,\ w/<4,s>}}]]$

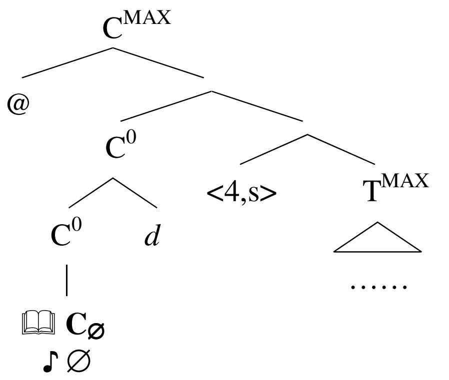
$= \lambda e\in D_E.\ \text{KNOW}(e) \wedge \forall d\in D_c[\text{d is compatible with what is known in e}$
$\quad \rightarrow \forall w\in D_s\ [\text{w is compatible with what is known in e}$
$\quad\quad \rightarrow [\![\mathbf{T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>,\ w/<4,s>}}]]$

$= \lambda e\in D_E.\ \text{KNOW}(e) \wedge \forall d\in D_c\ \forall w\in D_s\ [\text{d and w are compatible with what is known in e}$
$\quad \rightarrow [\![\mathbf{T^{MAX}}]\!]^{d,a^{e/<6,E>,\ d/<5,c>,\ w/<4,s>}}]$

This is actually the desired denotation of the computed node (See the next section to see how this would fit in the semantics of the whole sentence).

As we have just seen, attitude verbs quantify over indexical assignments under which their clausal complement is evaluated, and also give a restriction of the quantification over possible worlds for their clausal complement (see (11)). Then, since the matrix clause is by definition not the complement of an attitude verb, one might wonder how the semantics of the matrix clause is computed. An appropriate indexical assignment should be determined by the context in which the sentence is uttered. Also, the restriction of the quantification over possible worlds is the context set as we discussed earlier, and this is also determined by the discourse. If we assume that these are represented in the syntactic tree, we can obtain the desired semantics for the matrix clause. (12) illustrates the relevant tree.

(12)



Here, 🕮 $\mathbf{C}_\varnothing$ is the meaning of the phonetically empty matrix complementizer, and it is actually identical to 🕮 **that**. *d* represents the matrix indexical assignment and @ the (matrix) context set. *d* could be considered as a V head which is part of the head complex formed in the beginning of the derivation. @ may also be some head. Since we do not know much of the syntactic nature of d and @, I do not go into the syntactic details of these and leave this issue for future research. Also, I intend *d* and @ to be the denotations of their respective responsible syntactic objects, rather than the syntactic objects themselves. Thus, *d* is an element in D_c and @ an element in D_{st}. Nonetheless, I am using these in the syntactic tree for the sake of convenience. Now, let us see how the semantics of (12) is computed:

$$[\![\mathbf{C}^{\mathbf{MAX}}]\!]^{c,a}$$

$$= [\![@\ \mathbf{C}_\varnothing\ d\ \mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{c,a}$$

$$= [\![\mathbf{C}_\varnothing\ d\ \mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{c,a}(@) \quad \text{(by IFA)}$$

$$= [\![\mathbf{C}_\varnothing\ d\ ]\!]^{c,a}(\lambda g\in D_c.\ [\![\mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{g,a})(@)$$

$$= [\![\mathbf{C}_\varnothing\ ]\!]^{c,a}(d)(\lambda g\in D_c.\ [\![\mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{g,a})(@)$$

$$= [\lambda d'\in D_c.\ [\lambda f\in D_{<c,st>}.\ [\lambda u\in D_{st}.\ \forall w\in D_s\ [u(w) \rightarrow f(d')(w)]]]]$$
$$(d)(\lambda g\in D_c.\ [\![\mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{g,a})(@)$$

$$= \forall w\in D_s\ [@(w) \rightarrow [\lambda g\in D_c.\ [\![\mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{g,a}](d)(w)]$$

$$= \forall w\in D_s\ [@(w) \rightarrow [\![\mathbf{<4,s>}\ \mathbf{T}^{\mathbf{MAX}}]\!]^{d,a}(w)]$$

$$= \forall w\in D_s\ [@(w) \rightarrow [\lambda w'\in D_s.\ [\![\mathbf{T}^{\mathbf{MAX}}]\!]^{d,a^{w'/<4,s>}}](w)]$$

$$= \forall w\in D_s\ [@(w) \rightarrow [\![\mathbf{T}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>}}]$$

We can see that this is essentially equivalent to (1) above.

3.2 Situation as triplets of a world, a time and a space

In this section, I propose a decomposition of the situation into a product of worlds, times and spaces, and develop a semantic model that fits my derivation model that naturally deals with this decomposition.

According to Krazter (2006), situations have been made use of to evaluate linguistics expressions with respect to partial, rather than complete worlds. Situations appear to be useful in analyzing such linguistic phenomena as implicit domain restrictions and donkey sentences among others. In Soames' (200?) example in (13a), which describes a situation of a sleep lab, the domain for *everyone* should not contain the research assistants, since the research assistants should be awake so as to monitor the sleeping subjects. Kratzer explains that this can be accounted for if each predicate in one sentence comes with its own resource situation. On this approach, (13a) is analyzed as something like (13b).

(13) a. Everyone is asleep and is being monitored by a research assistant.

b. $\lambda s\ \forall x\ [[\text{person}(x)(s') \ \&\ s' \leq_P s] \rightarrow [\text{asleep}(x)(s)\ \&$
$\exists y[\text{research-assistant}(y)(s)\ \&\ \text{monitoring}(x)(y)(s)]]$

Now, consider the sentences in (14), which describe Donkey Parade:

(14) a. Whenever a donkey appeared, it was greeted enthusiastically.

b. Whenever a donkey appeared, the donkey was greeted enthusiastically.

The pronoun *it* in (14a) is an E-type pronoun which is interpreted like the corresponding definite description in (14b). Kratzer argues that *whenever* quantifies over minimal subsituations of Donkey Parade in which a donkey appeared. Thus, each of such situations contains exactly one donkey in it, and so the definite description *the donkey* is felicitous. Kratzer then gives the following analysis for the sentences in (14):

(15) $\lambda s\ \forall s'\ [[s' \leq_P s\ \&\ s' \in \text{Min}(\lambda s\ \exists x[\text{donky}(x)(s)\ \&\ \text{appeared}(x)(s)])]$
$\rightarrow \exists s''[s' \leq_P s''\ \&\ \text{greeted-enthusiastically}(\iota x\ \text{donkey}(x)(s'))(s'')]]$

Thus, analyses making use of situations are able to capture nicely the meaning of sentences involving implicit domain restrictions and of donkey sentences.

In what follows, I adopt this idea that each predicate is equipped with a situation argument that functions as a resource domain. To begin with, I'd like to clarify the notion of situation and to define it formally.

As Kratzer puts it, situations are sometimes taken to mean parts of a possible world. However, not only is this vague, it is even nonsensical. Formally speaking, a possible world is an element of D_s, i.e., an object that yields an extension when fed to a given intension. A possible world, by itself, then, is not something we conceive of when one talks about a "world". Thus, while one can talk about parts of a world in a casual context, possible worlds are formal objects that one cannot make sense in speaking parts of.

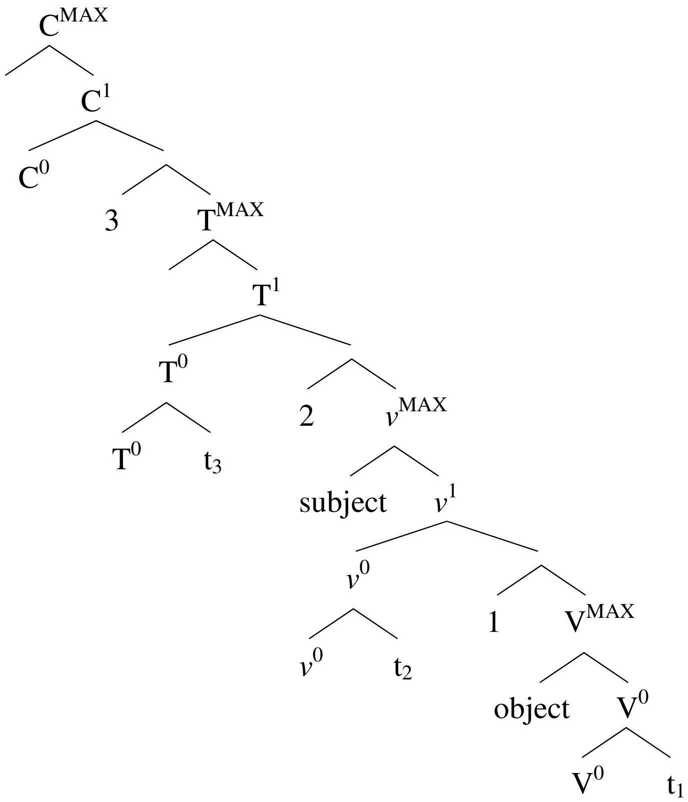
It seems to me quite obvious that in most cases where one talks about situations—that is, as long as we talk about the physical world—, one is considering some spatiotemporal regions coupled with some possible world for evaluation. For example, in the donkey sentence examples in (14), one is looking at all the (somehow defined) minimal spatiotemporal regions within the spatiotemporal extension of Donkey Parade in which exactly one donkey appeared in the actual world. Therefore, I define a situation as a triplet of a world, a time and a space. A possible world is an element of D_s. A time here is understood to be a time interval, an element in D_i. I assume that the space of a given situation is a subspace of some topological space, which may well be a three-dimensional Euclidean space $\mathbf{R}^3$ for most sentences. I call the type of such spaces p, so they are elements in D_p. Given these, situations are understood as elements of the Cartesian product of D_s, D_i and D_p, viz., $\{(w, t, s) \mid w \in D_s \wedge t \in D_i \wedge s \in D_p\}$.

Now, I discuss that situations enter into the semantics quite naturally in my model. To begin with, I assume that the meaning of a sentence generally involves quantification over events. For example, the sentence in (16) means that there is an event of John kissing Mary:

(16) John kissed Mary.

Therefore, the head that functions as the main predicate of a sentence takes an event argument. To implement this in my model, I propose that the *v* head that undergoes head movement is a quantifier over events, and it leaves behind a trace that functions as an event variable. Thus, in the following tree taken from Section 2, t_1 functions as an event variable:

(17)



Then, the transitive verb *kiss* has the following lexical entry:

(18) $[\![\textbf{kiss}]\!]^{c,a} = \lambda e \in D_E.\ [\lambda x \in D_e.\ \text{Theme}(e)(x) \wedge \text{KISS}(e)]$

Here, E is the type of events, and the Theme functions says that the theme of the event of e is x. Since whether or not certain events obtain depends on the evaluation world, *v* or V, which deals with events, needs access to the evaluation world. In my model, the world variable for the evaluation world of a clause is introduced as the trace of the head movement of the complementizer of the clause (see the previous section). Since this happens above the *v*P region, the introduced world variable (t_3 in (17)) is outside the *v*P. Then, the value of t_3 must be somehow passed downstairs in order that the evaluation world may be available in the *v*P region. The T head that is the sister of t_3 takes the denotation of t_3 and yields the denotation of the T head that dominates it and t_3. This non-terminal T head quantifies over times, so has undergone head movement and left behind the trace t_2, which is bound by binder 2. Since the moved head quantifies over times, what binder 2 abstracts over should involve a time, but if binder 2 abstracts over mere times, the value of t_3, namely, the evaluation world, cannot be passed down to the *v*P region. Therefore, I propose that what binder 2 abstracts over are pairs of the value of t_3 and the time that is quantified by the moved T head. In other words, binder 2 abstracts over elements of $D_{s\times i}$.

Since I have I have now introduced all the basic semantic types used in this paper and since we are talking about Cartesian products of semantics domains, let me list the basic semantic types for convenience and define semantic types for Cartesian products of semantic domains:

(19) Semantic domains

D_t := the set of truth values = {0,1}

D_e := the set of all individuals

D_E := the set of all events

D_p := the set of all spaces

D_i := the set of all times

D_s := the set of all possible worlds

D_c := the set of all indexical assignments

(20) For any semantic types σ, τ and υ,

$D_{\sigma\times\tau} := D_\sigma \times D_\tau := \{(x,y) \mid x \in D_\sigma \wedge y \in D_\tau\}$

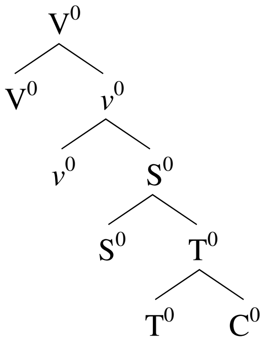
$D_{\sigma\times\tau\times\upsilon} := D_{\sigma\times\tau} \times D_\upsilon$

What I discussed in the previous paragraph is, then, that if a head quantifies over elements of type σ and takes an argument of type τ, it is possible for the binder associated with the head to abstract over variables of type $\sigma\times\tau$.

How, then, can spaces enter into our semantics so that situations may be constructed? It could be achieved easily if we either modified the meaning of C^0 so that it would quantify over world-space pairs or modified the meaning of T^0 so that it would quantify over time-space pairs. In such cases, binder 2 would automatically abstract over elements of $D_{s\times i\times p}$, i.e., situations. However, we do not have any strong reason to think that C^0 or T^0 quantifies over such pairs.

There is actually another way to construct situations in a natural way. If we assume that there is a head S^0 in the head complex that is responsible for quantification over spaces, then since it will be the case that we have heads that quantify over worlds, over times and over spaces respectively, it will be possible to constructed an abstraction over element in the Cartesian product of D_s, and D_i and D_p, i.e., situations. This in effect achieves quantification over situations. I adopt this approach for the rest of the paper, but the question as to whether or not such a head exists is not settled and remains as an open question. Furthermore, I assume that the SP is located below TP and above *v*P. Then, the derivation of a simple sentence starts with the formation of the head complex in (21) instead of the one we assumed in Section 2. S^0 is merged after T^0, because S^0 ends up being lower than T^0 when the derivation is complete.

(21)



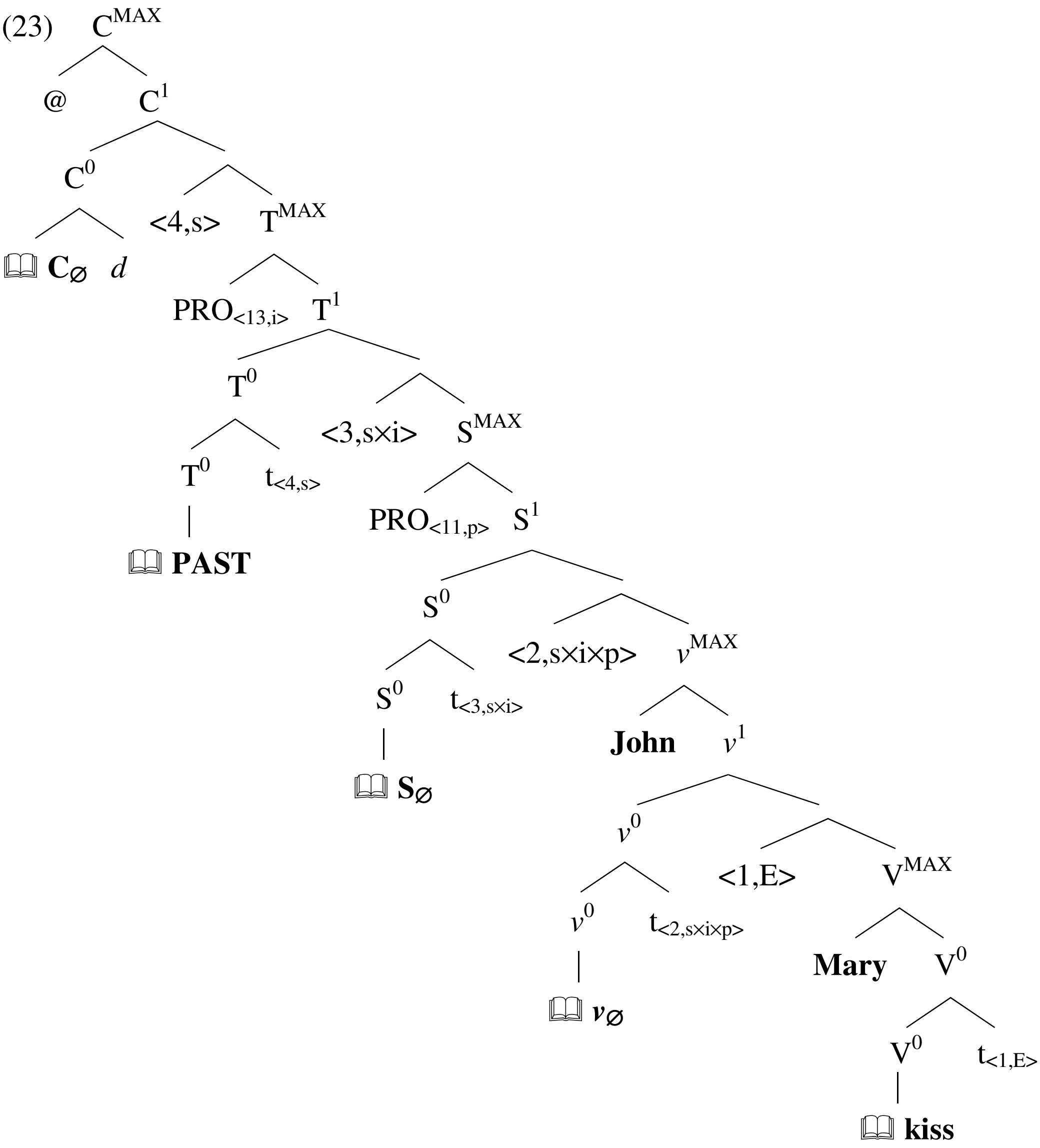
Before showing the LF after the derivation, I'd like to add a couple more of assumptions. Since T^0 and S^0 are quantifiers, it would be natural for them to have a domain restriction. Partee (1973) argues that the sentence in (22) does not mean either that there exists some time in the past at which I did not turn off the stove or that there exists no time in the past at which I turned off the stove, if uttered when halfway down the turnpike:

(22) I didn't turn off the stove.

Instead, the sentence refers to some particular contextually salient time, for example, the time of my getting ready to leave the house. I take such a time to be the restrictor to the temporal quantifier dented by T^0, and in order to capture this idea, I assume that there is a phonetically silent pronoun (represented as PRO) that denotes a time interval sitting at a specifier of TP and acts as a resource domain to the temporal quantifier. This temporal PRO has an index, and its value is given by an assignment provided by the context. In the same spirit, I also assume that SP has a spatial PRO at a specifier that denotes a space and provides a domain restriction to the spatial quantifier denoted by S^0. Again, the value of the PRO is given by an assignment.

Now, with all the assumptions introduced, the LF of the sentence in (16) looks like the following when the derivation is complete (actually, the subject John must have moved to a SpecTP position by A movement, but I have omitted it for the sake of simplicity since this movement has no semantic effect):

(23)



As can be seen from the tree, the terminal *v* head takes the trace of the head movement of S^0, and this trace function as a situation variable. Thus, v^0 quantifies over events that are situated in this situation. The following is the lexical entry of v^0:

(24) $[\![\boldsymbol{v_{\varnothing}}]\!]^{c,a} = \lambda\pi\in D_{s\times i\times p}.\ [\lambda f\in D_{Et}.\ [\lambda y\in D_e.$
$\exists e\in D_E$ [e occurs in $\pi \wedge$ Agent(e)(y) $\wedge$ f(e)]]]

Here, "occur" says that the event e is located in the situation π, an element in $D_{s\times i\times p}$. For the terminal T head and the terminal S head, the following lexical entries work:

(25) a. $[\![\textbf{PAST}]\!]^{c,a} = \lambda w \in D_s.\ [\lambda f \in D_{<s\times i,t>}.\ [\lambda t' \in D_i.$
$\exists t \in D_i\ [t < c(NOW) \wedge t \subseteq t' \wedge f((w,t))]]]$

b. $[\![\textbf{S}_{\varnothing}]\!]^{c,a} = \lambda \tau \in D_{s\times i}.\ [\lambda f \in D_{<s\times i\times p,t>}.\ [\lambda s' \in D_p.\ \exists s \in D_p\ [s \subseteq s' \wedge f((\tau,s))]]]$

$t \subseteq t'$ means that time interval t is included in time interval t', and $s \subseteq s'$ means that space s is included in space s'. The Agent function says that the agent of the event of e is y. As can be observed, these pair their first argument and a variable that their quantifier abstracts over to construct the argument to be fed to the function taken as their second argument (viz., f in their denotations). This is how their first argument is passed downstairs and the binder associated with them abstracts over elements in the corresponding Cartesian products.

Let us now compute the meaning of the sentence:

$[\![\textbf{C}^{\textbf{MAX}}]\!]^{c,a} = 1$

iff $[\![\textbf{C}_{\varnothing}]\!]^{c,a}\ (d)(\lambda d' \in D_c.\ [\![<4,s>\ \textbf{TP}]\!]^{d',a})(@) = 1$

iff $[\lambda d \in D_c.\ [\lambda f \in D_{<c,st>}.\ [\lambda u \in D_{st}.\ \forall w \in D_s\ [u(w) \rightarrow f(d)(w)]]]](d)$
$(\lambda d \in D_c.\ [\![<4,s>\ \textbf{TP}]\!]^{d',a})(@) = 1$

iff $\forall w \in D_s\ [@(w) \rightarrow [\lambda d' \in D_c.\ [\![<4,s>\ \textbf{T}^{\textbf{MAX}}]\!]^{d',a}](d)(w)] = 1$

iff $\forall w \in D_s\ [@(w) \rightarrow [\![<4,s>\ \textbf{T}^{\textbf{MAX}}]\!]^{d,a}(w)] = 1$

iff $\forall w \in D_s\ [@(w) \rightarrow [\lambda w' \in D_s.\ [\![\textbf{T}^{\textbf{MAX}}]\!]^{d,a^{w'/<4,s>}}](w)] = 1$

iff $\forall w \in D_s\ [@(w) \rightarrow [\![\textbf{T}^{\textbf{MAX}}]\!]^{d,a^{w/<4,s>}}] = 1$

Here,

$[\![\textbf{T}^{\textbf{MAX}}]\!]^{d,a^{w/<3,s>}} = 1$

iff $[\![\textbf{PAST}]\!]^{d,a^{w/<4,s>}}([\![t_{<4,s>}]\!]^{d,a^{w/<4,s>}})([\![<3,s\times i>\ \textbf{S}^{\textbf{MAX}}]\!]^{d,a^{w/<4,s>}})$
$([\![PRO_{<13,i>}]\!]^{d,a^{w/<4,s>}}) = 1$

iff $[\lambda w' \in D_s.\ [\lambda f \in D_{<s\times i,t>}.\ [\lambda t' \in D_i.\ \exists t \in D_i\ [t < d(NOW) \wedge t \subseteq t' \wedge f((w',t))]]]](w)$
$([\![<3,s\times i>\ \textbf{S}^{\textbf{MAX}}]\!]^{d,a^{w/<4,s>}})(a(<13,i>)) = 1$

iff $\exists t \in D_i\ [t < d(NOW) \wedge t \subseteq a(<13,i>) \wedge [\![<3,s\times i>\ \textbf{S}^{\textbf{MAX}}]\!]^{d,a^{w/<4,s>}}((w,t))]$
$= 1$

iff $\exists t \in D_i$ [$t < d(\text{NOW}) \wedge t \subseteq a(<13,i>)$
$\wedge [\lambda\tau \in D_{s\times i}.\ [\![\mathbf{S}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ \tau/<3,s\times i>}}]((w,t))] = 1$

iff $\exists t \in D_i$ [$t < d(\text{NOW}) \wedge t \subseteq a(<13,i>) \wedge [\![\mathbf{S}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}}] = 1$

Here,

$[\![\mathbf{S}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}} = 1$

iff $[\![\mathbf{S}_{\varnothing}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}}([\![t_{<3,s\times i>}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}})$
$([\![<2,s\times i\times p>\ \boldsymbol{v}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}})([\![\text{PRO}_{<11,p>}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}}) = 1$

iff $[\lambda\tau \in D_{s\times i}.\ [\lambda f \in D_{<s\times i\times p,t>}.\ [\lambda s' \in D_p.\ \exists s \in D_p\ [s \subseteq s' \wedge f((\tau,s))]]]]((w,t))$
$([\![<2,s\times i\times p>\ \boldsymbol{v}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}})(a(<11,p>)) = 1$

iff $\exists s \in D_p\ [s \subseteq a(<11,p>) \wedge [\![<2,s\times i\times p>\ \boldsymbol{v}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>}}((w,t,s))] = 1$

iff $\exists s \in D_p\ [s \subseteq a(<11,p>)$
$\wedge [\lambda\pi \in D_{s\times i\times p}.\ [\![\boldsymbol{v}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ \pi/<2,s\times i\times p>}}]((w,t,s))] = 1$

iff $\exists s \in D_p\ [s \subseteq a(<11,p>) \wedge [\![\boldsymbol{v}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>}}] = 1$

Here, I am writing (w,t,s) to mean ((w,t),s) for simplicity's sake, since this would cause no confusion.

$[\![\boldsymbol{v}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,\ s\times i\times p>}} = 1$

iff $[\![\boldsymbol{v}_{\varnothing}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>}}([\![t_{<2,s\times i\times p>}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,(w,t,s)/<2,s\times i\times p>}})$
$([\![<1,E>\ \mathbf{V}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>}})$
$([\![\mathbf{John}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>}}) = 1$

iff $[\lambda\pi \in D_{s\times i\times p}.\ [\lambda f \in D_{Et}.\ [\lambda y \in D_e.\ \exists e \in D_E\ [e \text{ occurs in } \pi \wedge \text{Agent}(e)(y) \wedge f(e)]]]]$
$((w,t,s))([\![<1,E>\ \mathbf{V}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>}})(\text{John}) = 1$

iff $\exists e \in D_E\ [e \text{ occurs in } (w,t,s) \wedge \text{Agent}(e)(\text{John})$
$\wedge [\![<1,E>\ \mathbf{V}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>}}(e)] = 1$

iff $\exists e \in D_E\ [e \text{ occurs in } (w,t,s) \wedge \text{Agent}(e)(\text{John})$
$\wedge [\lambda e' \in D_E.\ [\![\mathbf{V}^{\mathbf{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>,\ e'/<1,E>}}](e)] = 1$

iff $\exists e \in D_E$ [e occurs in (w,t,s) $\wedge$ Agent(e)(John)
$\wedge$ $[\![\mathbf{V^{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>,\ e/<1,E>}}$] = 1

Finally,

$$[\![\mathbf{V^{MAX}}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>,\ e/<1,E>}} = 1$$

iff $[\![\textbf{kiss}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>,\ e/<1,E>}}$
$([\![t_{<1,E>}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>,\ e/<1,E>}})$
$([\![\textbf{Mary}]\!]^{d,a^{w/<4,s>,\ (w,t)/<3,s\times i>,\ (w,t,s)/<2,s\times i\times p>,\ e/<1,E>}}) = 1$

iff $[\lambda e' \in D_E.\ [\lambda x \in D_e.\ \text{Theme}(e')(x) \wedge \text{KISS}(e')]](e)(\text{Mary}) = 1$

iff $\text{Theme}(e)(\text{Mary}) \wedge \text{KISS}(e) = 1$

In sum, (16) has the truth condition in (26a), and (26b) is an English translation of it:

(26) a. $\forall w \in D_s$ [@(w) $\rightarrow$
$\exists t \in D_i$ [t < d(NOW) $\wedge$ t $\subseteq$ a(<13,i>)
$\wedge$ $\exists s \in D_p$ [s $\subseteq$ a(<11,p>)
$\wedge$ $\exists e \in D_E$ [e occurs in (w,t,s) $\wedge$ Agent(e)(John) $\wedge$ Theme(e)(Mary)
$\wedge$ KISS(e)]]]]

b. For all w such that @(w), there is a time t such that t < d(NOW) and t $\subseteq$ a(<13,i>) and there is a space s such that s $\subseteq$ a(<11,p>), and there is an event e such that the agent of e is John and the theme of e is Mary and e is a kissing event that occurs in the situation specified by (w,t,s).

3.3. Dependence of noun phrase interpretation

In this section, I review Musan's observation of the temporal interpretation of noun phrases and then look at more data to argue that similar observation holds for worlds and spaces.

It has been known that the temporal interpretation of a noun phrase can be independent of the temporal interpretation of the main predicate of the clause in which it appears. For example, in (27), taken from Musan (1995), the person who was hiding might have already murdered John, and so was a murderer at the same time of the hiding. However, the sentence can also be true in a scenario in which he has murdered John by now but had not yet murdered John when he hid behind the curtain. It is even possible that he has not yet committed the murder but is supposed to kill John in the future.

(27) John's murderer hid behind the curtain.

Musan (1995) observes, however, that only presuppositional noun phrases can be temporally independent of the interpretation of the main predicate and that cardinal noun phrases always receive temporally dependent interpretations. For convenience, I refer to this as Musan's Thesis of Temporal Dependence. *John's murderer* in (27) is a definite description, and so it is presuppositional. That is why *John's murderer* can have temporally independent interpretations. Musan's point is illustrated in the following pairs of sentences.

(28) a. In the forties, all professors were young.

b. In the forties, professors were young.

(29) a. Last year, some of the congressmen came to the party.

b. Last year, some congressmen came to the party.

In (28a) and (29a), *all professors* and *some of the congressmen* can be temporally independent of the main predicate of the clauses. That is, these can be interpreted as "all people who are professors now" and "some of the people who are congressmen now" respectively. This is because these noun phrases are presuppositional: in (28a), *all* is a strong determiner, and in (29a), the noun phrase is a partitive. On the other hand, in (28b) and (29b) (on its most natural reading, where *some congressmen* is interpreted as cardinal), the subject noun phrases can only be temporally dependent.

Now, I argue that Musan's observation on the distribution of temporally dependent and independent interpretations of noun phrases is generalized and that parallel generalizations hold of worlds and spaces. In the last chapter of Musan (1995), she briefly discusses such possibilities.

For worlds, what is expected would be that the world interpretation of cardinal noun phrases is dependent on the world interpretation of the main predicate, but the world interpretation of presuppositional noun phrases can be independent (henceforth the Thesis of World Dependence). Musan gives the following pair as initial evidence for this ((30a) is taken from Abusch (1994, 104) and (30b) is due to Kai von Fintel (p.c.)):

(30) a. Things would be different if <u>a senator</u> had grown up to be a rancher instead.

b. Things would be different if there was <u>a senator</u> having grown up to be a rancher instead.

For (30a), where *a senator* could be presuppositional, both the readings roughly described in (31) obtain. By contrast, for (30b), where *there* forces *a sanotor* to be cardinal, only the reading in (31a) obtains.

(31) a. For all worlds w such that there is a senator <u>in w</u> that grew up to be a rancher in w, things are different in w.

b. For all worlds w such that there is a senator in the actual world that grew up to be a rancher in w, things are different in w.

I think that Musan's conjecture is actually right. For more support, let us look at some more examples. Consider (32):

(32) John thinks that Mary kissed a student.

First, suppose that we have not been talking about any students in the discourse. Then, *a student* in (32) will have a cardinal interpretation. In this case, (33a) seems to be the only available reading for (32). On the other hand, if *a student* in (32) is presuppositional (i.e., if it is understood as "one of the students"), then both the readings in (33) obtain[1].

(33) a. For every world w that is compatible with what John believes in the actual world w*, there is an individual x such that x is a student in w and Mary kissed x in w.

b. There is an individual x such that x is a student in the actual world w* such that for every world w that is compatible with what John believes in w*, Mary kissed x in w.

The same point is made for the following sentence as well:

(34) Mary must kiss a student.

Again, on the reading where *a student* is cardinal, of the two interpretations roughly sketched in (35), (35a) is the only available one. By contrast, on the reading where *a student* is presuppositional, both (35a) and (35b) obtain.

(35) a. For every regulation world w accessible from the actual world w*, there is an individual x such that x is a student in w and Mary kisses x in w.

b. There is an individual x such that x is a student in the actual world w* such that for every regulation world w accessible from the actual world w*, Mary kisses x in w.

These data all support the Thesis of World Dependence. One might think that intensional verbs that take a noun phrase argument might give counterexamples to the thesis. Consider (36):

1. For the sake of simplicity, I ignore the non specific *de re* reading in the discussion here, but see subsection 3.4.3 for a relevant discussion.

(36) a. Mary wants a chinchilla.

b. Mary imagines a chinchilla.

c. Mary is looking for a chinchilla.

In (36), even on the readings where *a chinchilla* is not presuppositional, its world interpretation appears to be independent, since the evaluation world for *a chinchilla* is not the actual world. However, according to the so-called sentensionalist or propositionalist approach, these intensional verbs that do not appear to take a clausal complement actually have a clausal complement that is somehow hidden on the surface, and the source of the apparent intensionality of the argument noun phrase is due to this clausal complement (see Den Dikken, Larson and Ludlow (1997)). If we take this line, then the world interpretation of *a chinchilla* in (36) would actually be dependent on the predicate of the hidden clausal complement of the intensional verb. Therefore, the Thesis of World Dependence can be maintained.

Finally, let us look at the spatial interpretation of noun phrases. Musan suggests investigating the possibility of distinguishing spatially dependent and spatially independently noun phrases as well. Although Musan does not give any sentence regarding this, I think that the same generalization holds of spaces. That is, the spatial interpretation of cardinal noun phrases is dependent on the spatial interpretation of the main predicate, but the spatial interpretation of presuppositional noun phrases can be independent (henceforth the Thesis of Spatial Dependence).

Given Musan's Thesis of Temporal Dependence, the Thesis of Spatial Dependence is actually necessarily true, so long as we assume that one and the same individual may only be located in one place at a given time. Consider (37):

(37) a. Mary met a student.

b. Mary met the student.

On the cardinal reading of *a student* in (37a), the predication time of *student* must coincide with the time of meeting (say 3 o'clock), since *a student* is cardinal. In order for the individual predicated of by *student* to be involved in the event of Mary's meeting, at 3 o'clock, the individual obviously has to be situated in the same space as the meeting takes place in (say, in the garden). As we have just established, however, at 3 o'clock, the individual must be a student (in somewhere in the world). Since by assumption, one individual cannot be located in two different places, the garden is the only place that this individual is located in at 3 o'clock. Therefore, it is concluded that the individual that Mary met was a student at 3 o'clock and in the garden.

On the other hand, in (37b), *the student* can have a temporally independent interpretation, since it is presuppositional. For example, suppose the following scenario: John is the only student that Mary taught 5 years ago at college. He has already graduated and is no longer a student. Yesterday, John came to visit this famous garden for the first time, and Mary met him there. If the previous context has been referring to the unique student Mary taught 5 years ago, (37b) is felicitously uttered to mean Mary met John.

When the meeting takes place, John obviously has to be located in the garden, but since this was John's first visit to the garden, when John was a student, he was never in the garden. Therefore, the space in which John was a student does not coincide with the space in which the meeting took place.

If the Thesis of Spatial Dependence is a mere corollary of the Thesis of Temporal Dependence in conjunction with the plausible assumption that an individual cannot be simultaneously located in more than one place, one might question whether there is any point in positing the Thesis of Spatial Dependence at all. Here, I will not try to discuss exclusively whether or not the Thesis of Spatial Dependence should be independently motivated, since all I need is the generalization of the distribution of spatially dependent and spatially independent noun phrases. However, I would like to discuss some examples that might be taken as independent evidence for the Thesis of Spatial Dependence.

(38) is a dispositional sentence and means something like "whenever John is in the pub, he is talkative, and whenever he is outside, he is silent."

(38) John is a talkative man in the pub, but is a silent man outside.

If a dispositional sentence is evaluated with respect to a certain (somewhat long) time interval during which the described disposition obtains without necessarily having the time interval completely filled with events or situations that embody the disposition, the two conjuncts in (38) would be evaluated with respect to one and the same present time interval, and they do not have to be contradictory even though John can never be both in the pub and outside simultaneously. Thus, using dispositional sentences allows a possibility of having two predicates truthfully predicate of one individual for one and the same time interval, but for two distinct spaces. Now consider the following:

(39) a. In the garden, Mary met a talkative man yesterday.

b. In the garden, Mary met the talkative man yesterday.

Suppose a scenario in which Mary met John in the garden and he is as described in (38) above. Then, (39a) is false when *a talkative man* has a cardinal reading, since (a) entails that the man who Mary met was talkative in the garden. By contrast, even though John was silent in the garden, (39b) can be truthfully uttered, provided that John has been established to be a unique salient man who is talkative in certain occasions. The same point is illustrated by (40) and (41). Suppose that (40) is a true description of John. In a scenario in which Mary met John in the garden, (41a) should be wrong on its cardinal reading, but (41b) can be truthfully uttered, provided that the garden is not in the disco.

(40) John is a drag queen in that disco, but otherwise he isn't.

(41) a. In the garden, Mary met a drag queen.

b. In the garden, Mary met the drag queen.

Thus, in any case, the Thesis of Spatial Dependence seems to hold.

With the Theses of World Dependence, of Temporal Dependence and of Spatial Dependence established, we have the following generalization:

(42) The world, time, and space interpretations of cardinal noun phrases are dependent on the main predicate.

Since a world, a time and a space comprise a situation, this means that the situational interpretation of cardinal nouns is dependent of that of the main predicate. In the next section, I discuss how this can be derived in our system.

3.4 Presuppositionality and Binding Theory

In this section, in order to explain the difference between cardinal noun phrases and presuppositional noun phrases we have seen in the previous section, I propose that cardinal noun phrases have an anaphor that denotes a situation, whereas presuppositional noun phrases have quantifiers inside them just like a clause.

Musan presents a possible analysis of her Thesis of Temporal Dependence based on Diesing (1992). Diesing argues that presuppositional noun phrases (including noun phrases with a weak determiner in their strong meaning) are outside the VP domain (which would correspond to *v*P in the recent literature) at LF, whereas cardinal noun phrases remain inside the VP domain. Her mapping hypothesis states that material outside the VP at LF is mapped to the restrictive clause and material inside the VP domain at LF is mapped to the nuclear scope in the sense of discourse representation theory (Kamp (1981), Heim (1982)). Then, it seems that the temporal interpretations of presuppositional and cardinal noun phrases can be nicely accounted for if there is some temporal quantifier just above VP that obligatorily binds the time arguments of nouns that are inside VP together with the time argument of the verb. On this approach, cardinal noun phrases are necessarily temporally dependent, since they stay inside VP and so their time arguments are obligatorily bound by the temporal quantifier. On the other hand, presuppositional noun phrases get scoped out of this temporal quantifier, and so their time arguments cannot be bound by the quantifier, and hence they obtain temporally independent readings.

Musan argues, however, that this scope approach does not work since it cannot account for certain data. To account for the difference between presuppositional and cardinal noun phrases, she instead proposes that presuppositional noun phrases contain an extra quantification over stages of individuals, i.e., temporal slices of individuals. In Appendix, I discuss Musan's theory in detail and argue that the scope approach is actually tenable.

Following Musan, Kusumoto (2005) analyzes strong determiners as containing a quantifier over times, rather than over stages of individuals. For example, the denotation of *every* is as follows:

(43) $[\![\textbf{every}]\!]^g = \lambda P \in D_{<e,<i,st>>}. [\lambda Q \in D_{<e,<i,st>>}. [\lambda t \in D_i. [\lambda w \in D_s.$ for every individual x such that there is a time t' such that $P(x)(t')(w) = 1$, $Q(x)(t)(w) = 1]]]$

Kusumoto explains that (44) is judged true in a situation where there is a group of five people who were fugitives at different times in the past but are currently in jail. This fact is nicely captured with the denotation in (43).

(44) Every fugitive is in jail.

Since cardinal noun phrases have only temporally dependent interpretations, unlike Musan, Kusumoto proposes that they do not involve temporal quantifiers. Thus, the following is the denotation of *a* on its cardinal reading:

(45) $[\![\mathbf{a}]\!]^g = \lambda P \in D_{<e,<i,st>>}. [\lambda Q \in D_{<e,<i,st>>}. [\lambda t \in D_i. [\lambda w \in D_s.$ there is an individual x such that $P(x)(t)(w) = 1$ and $Q(x)(t)(w) = 1]]]$

As is obvious from the denotation, in order for a cardinal noun phrase to be interpreted properly, it is necessary that it combine with the main predicate before the main predicate's time argument is saturated.

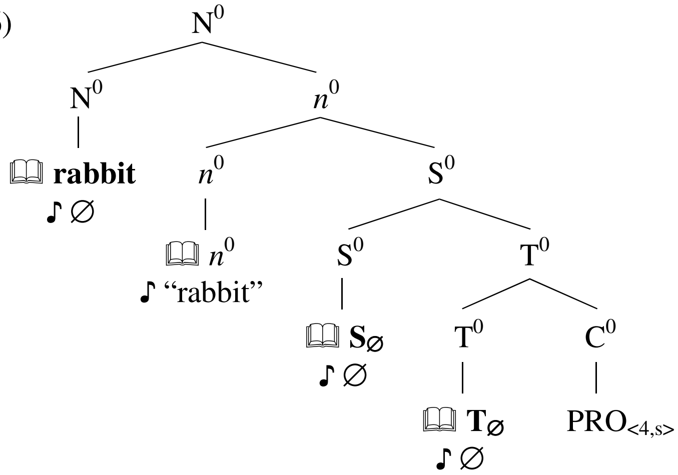
There are two important differences between Musan's theory and Kusumoto's theory (see Appendix for the details of Musan's theory). First, in Kusumoto's theory, the time argument of the main predicate is given by the phonetically null operator in the main clause which acts as a quantifier over times, and not via the quantifier inside a noun phrase taken by the predicate as in Musan's theory. Secondly, since unlike Musan's theory, the temporal interpretation of cardinal noun phrases is truly dependent on the temporal interpretation of the main predicate, it is predicted that the cardinal noun phrases would only appear below the phonetically null tense operator. This means that cardinal noun phrases should be inside the VP at LF as Diesing's theory predicts, provided that the tense operator is just above the VP. Thus, Kusumoto's approach could be regarded as a mixed theory that utilizes both scope and noun internal properties (see Appendix for a relevant discussion).

In what follows, I basically adopt Kusumoto's approach to presuppositional noun phrases, but offer a different analysis of cardinal noun phrases.

3.4.1 Presuppositional noun phrases

I follow Kusumoto in that presuppositional noun phrases involve a quantification over times, and in order to account for the world independence and the spatial independence of presuppositional noun phrases, I also propose that presuppositional noun phrases involve quantifications over worlds and over spaces as well. Then, the most natural assumption is that a noun phrase is derived from a head complex containing S^0, T^0 and C^0 just like the derivation of a clause. Also, I posit n^0 as the nominal counterpart of v^0. Like v^0, n^0 quantifies over events. Then, a presuppositional noun phrase with *rabbit* starts with the head complex in (46) below:

(46)



Since the phonetic content of V^0 moves to v^0 in English, I assume by analogy that the phonetic content of N^0 moves to n^0. The terminal S^0 has the same meaning as in a clause. The denotation of the terminal T^0 here is also similar to the that of terminal T^0s found in a clause, but unlike in a clause, the tense of noun phrases does not seem to be restricted to the past, present, etc. (see (44)), I assume that the domain of the quantification of the T^0 in a noun phrase has no such restriction. Thus, the following is the lexical entry for $T_{\varnothing}$:

(47) $[\![\mathbf{T}_{\varnothing}]\!]^{c,a} = \lambda w \in D_s. [\lambda f \in D_{<s\times i,t>}. [\lambda t' \in D_i. \exists t \in D_i [t \subseteq t' \wedge f((w,t))]]$

Note that the C^0 is not a quantifier, but a PRO that acts as a world variable. When the main predicate (i.e., the predicate of which the noun phrase is an argument of) is in the matrix clause, this PRO is necessarily co-indexed with the trace of head movement of the C^0 in the matrix clause of the sentence and thus gets bound by the world binder created by this head movement of C^0. For example, in (48), the presuppositional noun phrase *every rabbit* quantifies over actual rabbits. This is done by virtue of the PRO in the noun phrase *every rabbit* being bound by the world binder created by the matrix C^0.

(48) Every rabbit in the garden is cute.

On the other hand, when the main predicate is in an embedded clause, the PRO gets bound either by the world binder created by head movement of the C^0 in the clause of the main predicate (local binding) or by the world binder created by head movement of the C^0 in some higher clause (long-distance binding). Local binding results in opaque (or *de dicto*) readings, whereas long-distance binding gives rise to transparent (or *de re*) readings. For instance, consider (49):

(49) Mary thinks that every student in the garden is cute.

This sentence exhibits two readings roughly described in (50):

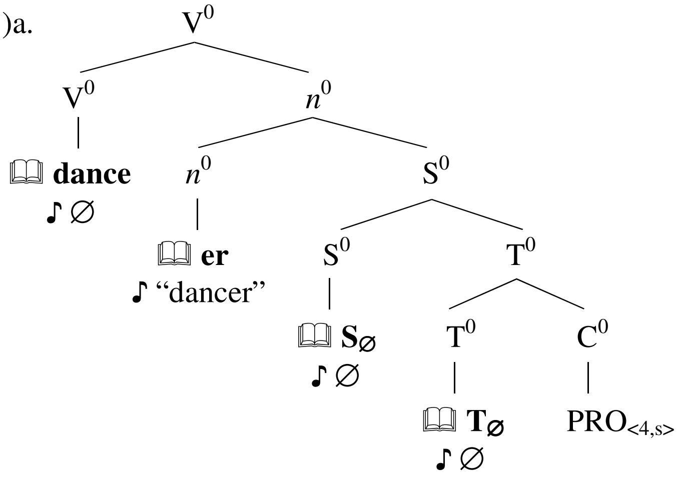
(50) a. In every world w that is consistent with what Mary thinks in the actual world w*, every student in the garden in w is cute in w.

b. In every world w that is consistent with what Mary thinks in the actual world w*, every student in the garden in w* is cute in w.

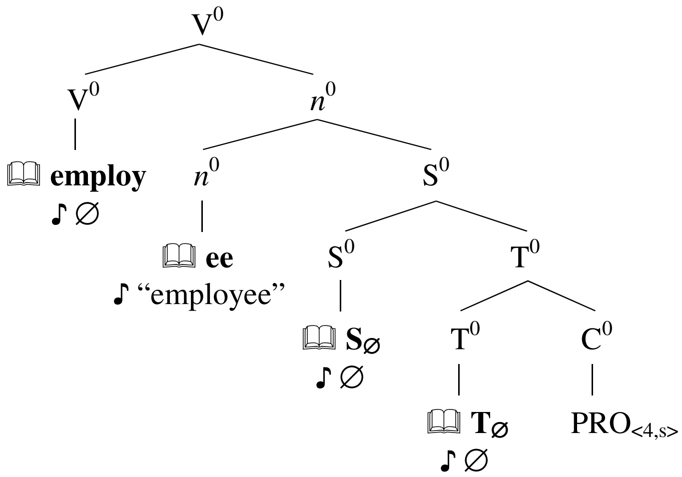
In (50a), *every student* has an opaque reading, and this is achieved by having the PRO in the noun phrase *every student* bound locally (i.e., by the binder created by the head movement of *that*). In (50b), *every student* has a transparent reading, and in this case, the PRO is bound long-distance (i.e., by the binder created by the head movement of the invisible matrix C^0). I do not know if the long-distance binding requires movement of the noun phrase to the matrix clause.

Now, let me explain that the postulation of n^0 is not without any plausible reasons. English has suffixes *-er* and *-ee*, which turn verbs into nouns. For any given verb *X*, *X-er* means "individual who X's" and *X-ee* means "individual who is X'ed". This phenomenon seems to be naturally captured if *-er* and *-ee* are in n^0. Recall that n^0 has been postulated as the nominal counterpart of v^0. The idea is that *-er* corresponds to the v^0 for the active voice and *-ee* corresponds to the v^0 for the passive voice. On this view, the derivation of *every dancer* and *every employee* start with the head complexes in (51a) and (51b) respectively:

(51)a.

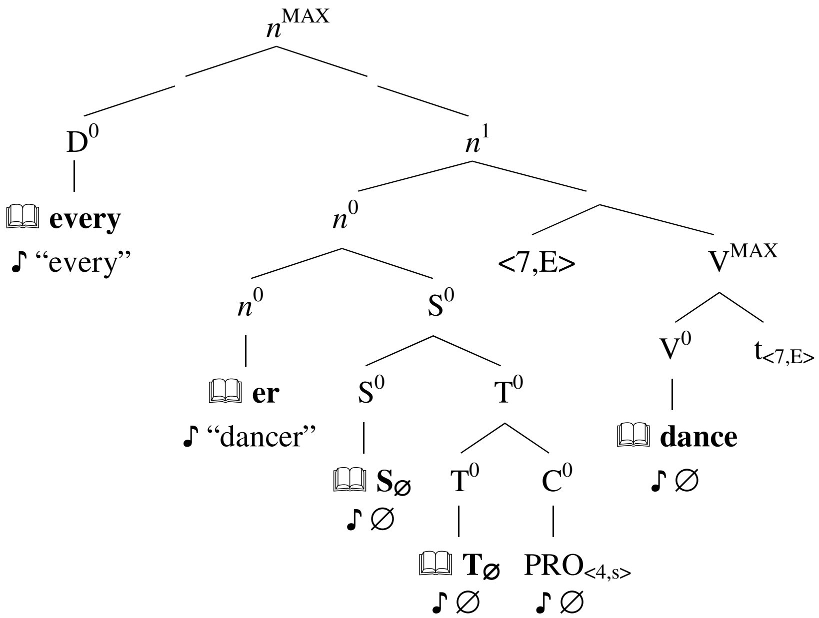


b.



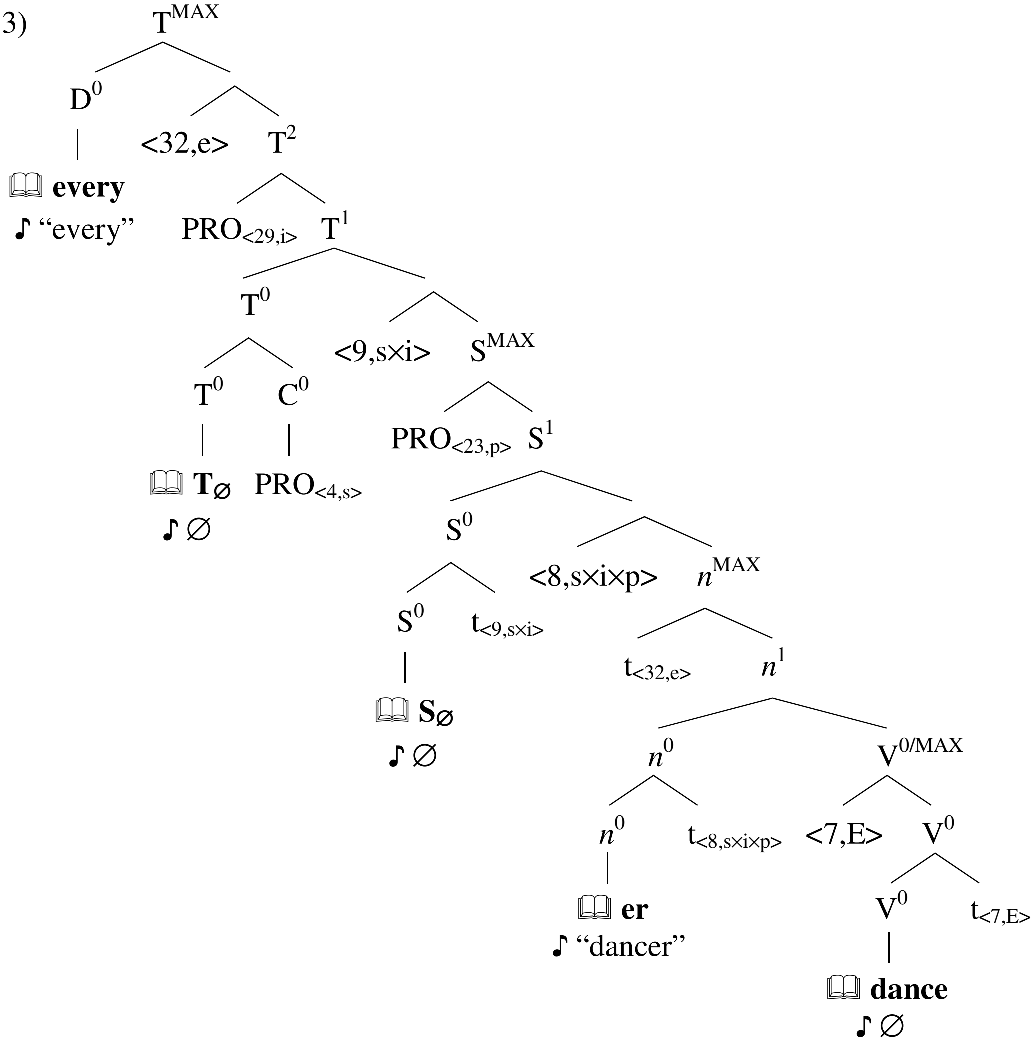
The derivation of a presuppositional noun phrase is parallel to that of a clause. While n^0 corresponds to v^0, determiners like *every* correspond to individual arguments (i.e., subjects and objects) of predicates in a clause. Since a dancer is an individual who is the agent of an event of dancing, the determiner in *every dancer* corresponds to the subject of the verb *dance*. Therefore, *every* is merged as a specifier of *n*P after the head movement of n^0:

(52)



Since S^0 and T^0 are quantifiers, they need to move. Thus, the derivation proceeds with head movement of S^0 followed by head movement of T^0. In order to obtain a proper interpretation, it is necessary that the determiner move to a SpecTP position:

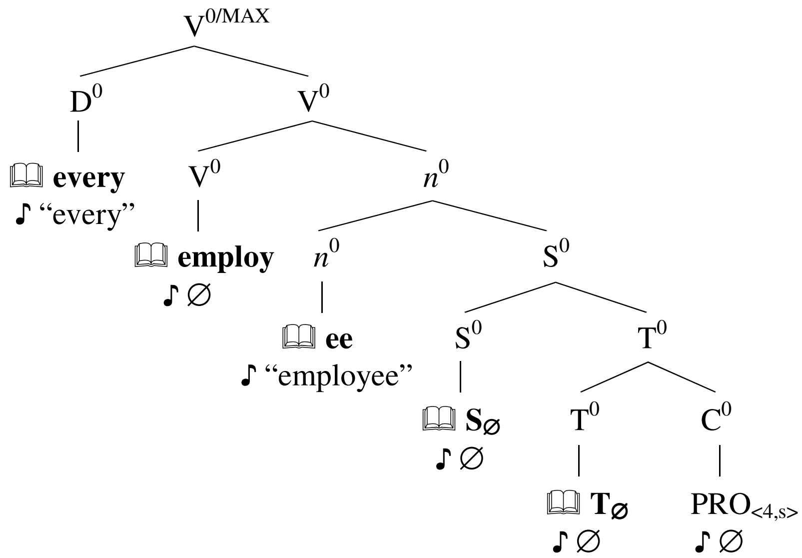
(53)



Again, the derivation looks parallel to the derivation of a clause, since in a clause, the subject moves to SpecTP (presumably by A movement). This movement might possibly be done via a SpecSP position. Note that C^0 does not undergo head movement, since it is not a quantifier.

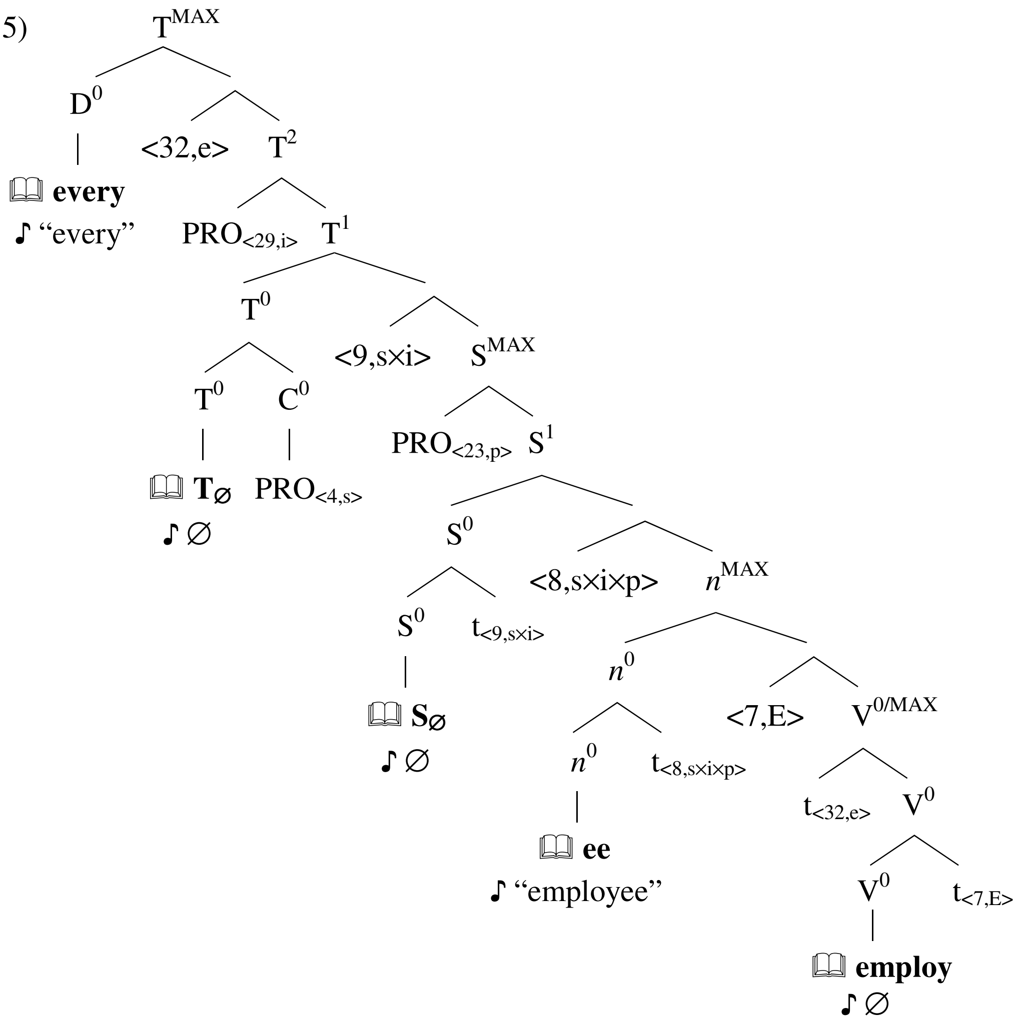
In the case of the derivation of *every employee*, since an employee is an individual who is the theme of an event of employment, the determiner *every* corresponds to the object of the verb *employ*. Therefore, *every* is directly merged with the head complex:

(54)



Then, as in the derivation of every dancer, head movement of n^0, S^0 and T^0 takes place. The determiner ends up at a SpecTP position, possibly via a Spec*n*P and a SpecSP on the way. This is parallel to passive formation in a clause.

(55)



As you can see, in my model, presuppositional noun phrases are TPs, rather than NPs. I refer to these TPs as noun phrases for convenience.

The lexical entry of *-er* should be the same as that of v^0 proposed in Section 3.2. The lexical entry of *-ee* is different from that *-er* in that it does not take an individual argument. Hence the following lexical entries:

(56) a. $[\![\mathbf{er}]\!]^{c,a} = \lambda\pi\in D_{s\times i\times p}.\ [\lambda f\in D_{Et}.\ [\lambda y\in D_e.\ \exists e\in D_E\ [\text{e occurs in } \pi \wedge \text{Agent(e)(y)} \wedge \text{f(e)}]]]$

b. $[\![\mathbf{ee}]\!]^{c,a} = \lambda\pi\in D_{s\times i\times p}.\ [\lambda f\in D_{Et}.\ \exists e\in D_E\ [\text{e occurs in } \pi \wedge \text{f(e)}]]$

Since *-ee* corresponds to the passive version of v^0, the passive version of v^0 ($v^{PASS}{}_{\varnothing}$) must have the same lexical entry as *-ee*:

(57) $[\![\boldsymbol{v}^{\mathbf{pass}}{}_{\varnothing}]\!]^{c,a} = \lambda\pi\in D_{s\times i\times p}.\ [\lambda f\in D_{Et}.\ \exists e\in D_E\ [\text{e occurs in } \pi \wedge \text{f(e)}]]$

Note that this means that we need the v^0 head even in passive sentences since we need quantification over events for passive sentences as well. Now, with the lexical entry of *every* below, we can compute the denotations of *every dancer* and *every employee* as in (59).

(58) $[\![\textbf{every}]\!]^{c,a} = \lambda f\in D_{et}.\ [\lambda g\in D_{et}.\ \forall x\ [f(x) \rightarrow g(x)]]$

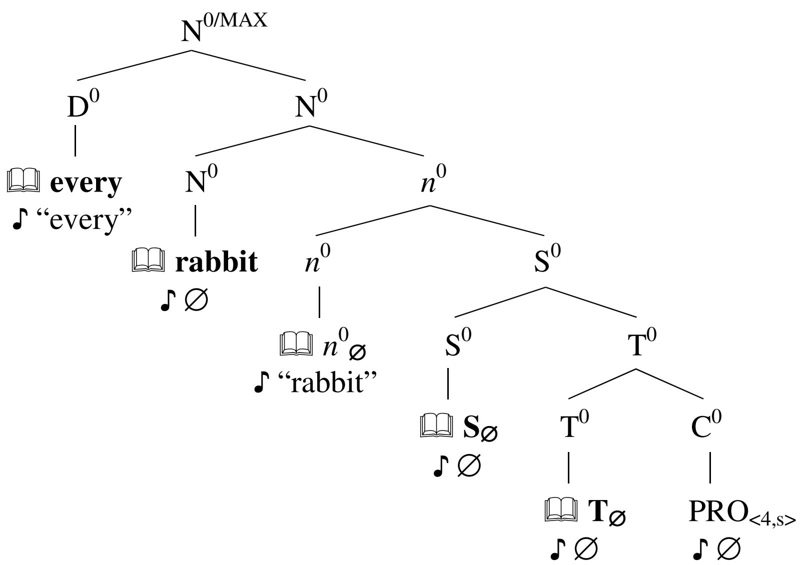
(59) a. $[\![\textbf{every dancer}]\!]^{c,a}$
$= \lambda g\in D_{et}.\ \forall x\ [\exists t\in D_i\ [t \subseteq a(<29,i>) \wedge \exists s\in D_p\ [s \subseteq a(<23,p>)$
$\wedge \exists e\in D_E\ [e \text{ occurs in } (a(<4,s>),t,s) \wedge \text{Agent}(e)(x) \wedge \text{DANCE}(e)]]$
$\rightarrow g(x)]$

b. $[\![\textbf{every employee}]\!]^{c,a}$
$= \lambda g\in D_{et}.\ \forall x\ [\exists t\in D_i\ [t \subseteq a(<29,i>) \wedge \exists s\in D_p\ [s \subseteq a(<23,p>)$
$\wedge \exists e\in D_E\ [e \text{ occurs in } (a(<4,s>),t,s) \wedge \text{Theme}(e)(x) \wedge \text{EMPLOY}(e)]]$
$\rightarrow g(x)]$

Note that the denotation of *every* itself does not have a domain restriction. Recall Kratzer's discussion we have looked at in Section 3.1 that the domain restriction of a noun phrase can be captured as a situation. In my model, this situation is given by virtue of the world PRO, the temporal quantifier (T^0) and the spatial quantifier (S^0) in the presuppositional noun phrase. Since these quantifiers come with their own domain restrictions that are given by resource domain PROs, the domain of the constructed situation variable gets restricted accordingly.

Now, let us return to *every rabbit*. In this case, the relevant event involved would be an (stative) event of some individual's being a rabbit. It seems to me that the individual who is a rabbit in such an event should be regarded as the Theme of the event rather than the Agent, since one cannot be a rabbit intentionally. Then, *rabbit* would be an "unaccusative" noun, so to speak, and so the determiner must be merged in the "object" position just as in the case of *every employee*:

(60)



Here, $n^0_\varnothing$ is similar to *-ee* in that it lacks an agent role argument. Let us now note that in order to say that an individual is a rabbit in a given situation, it seems necessary to require that the individual be a rabbit no matter where and when the individual is located in the situation. Thus, it would not be enough to just say that there is an event that occurs in a given situation. To capture the intuition, I utilize the materialization function M: $D_E \rightarrow D_{s\times i\times p}$, which, when given an event, yields the exact situational region in which the event physically occurs. Then, I propose the following lexical entry for $n^0_\varnothing$:

(61) $[\![\boldsymbol{n_\varnothing}]\!]^{c,a} = \lambda\pi\in D_{s\times i\times p}.\ [\lambda f\in D_{Et}.\ \exists e\in D_E\ [M(e) \supseteq \pi \wedge f(e)]]$

$M(e) \supseteq (a(<4,s>),t,s)$ means that the situational region in which e occurs is identical or includes the situation π. This kind of consideration is probably required in considering the denotations of *dancer* and *employee* and also of v^0, but since this appears to be a very complicated task, I leave it for future research, and content myself with the expression "e occurs in π" for most cases for now. The denotation of *every rabbit* is now calculated as follows (assuming that $PRO_{<23,p>}$ is at a SpecSP and $PRO_{<29,i>}$ is at a SpecTP):

(62) $[\![\textbf{every rabbit}]\!]^{c,a}$
$= \lambda g\in D_{et}.\ \forall x\ [\exists t\in D_i\ [t \subseteq a(<29,i>) \wedge \exists s\in D_p\ [s \subseteq a(<23,p>)$
$\wedge\ \exists e\in D_E\ [M(e) \supseteq (a(<4,s>),t,s) \wedge Theme(e)(x) \wedge\ RABBIT(e)]] \rightarrow g(x)]$

Here, RABBIT(e) means that e is an (stative) event of somebody's being a rabbit.

Lastly, let me briefly mention the definite article. I adopt Heim & Kratzer (1998)'s analysis of the definite article *the*:

(63) $[\![\textbf{the}]\!]^{c,a} = \lambda f\in D_{et}$: there is exactly one $x \in D_e$ such that $f(x) = 1$.
the unique $x \in D_e$ such that $f(x) = 1$

Unlike determiners like *every*, which have denotations of type <et,<et,t>>, the semantic type of the denotation of the is <et,e>. Nevertheless, I assume that presuppositional noun phrases with *the* are derived just the same way as the above cases except that *every* is replaced with *the*. Thus, *the* leaves behind a trace of type e. See section 4.2.3 for a relevant discussion.

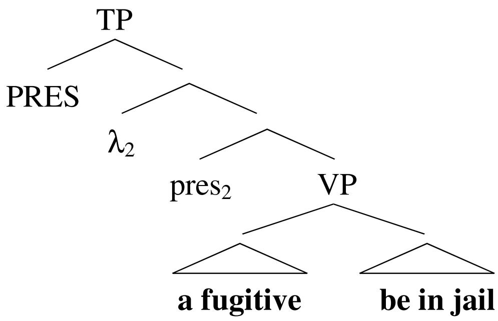
As we have seen above, it seems that it is whether there is a v^0 or n^0 in the head complex that determines whether the derived phrase is a verb phrase or noun phrase. Then, it may be the case that we do not need to distinguish V and N. I would like to leave this issue for future investigation.

3.4.2 Cardinal noun phrases

Although I have basically adopted Kusumoto's approach to presuppositional noun phrases, I differ from Kusumoto in explaining how cardinal noun phrases receive their temporal dependent interpretations. In this subsection, to account of why cardinal noun phrases' world, time and space interpretations are dependent on the respective interpretations of the main predicate, I propose that cardinal noun phrases have an situation anaphor inside them.

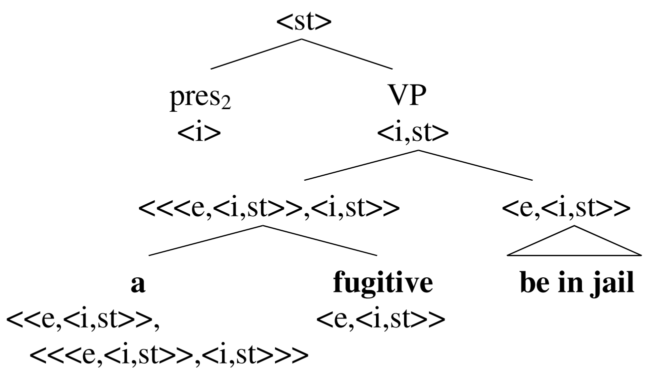
Kusumoto argues that unlike main predicates, noun phrases do not have a syntactically present time argument. Therefore, a cardinal noun phrase obtains its temporally dependent meaning by virtue of taking the denotation of the main predicate before the main predicate's time argument is saturated. Since the time argument of the main predicate is the tense morphology in Kusumoto's system, the tense morphology is merged after the cardinal noun phrase and the main predicate are merged. Thus, the sentence *there is a fugitive in the jail* has the following tree structure:

(64)



This way, the tense morphology in effect saturates the time argument of the noun and that of main predicate simultaneously, as is clear from the denotation in (45) and (65) below. This is Kusumoto's account of why the temporal interpretation of cardinal noun phrases is dependent of the temporal interpretation of the main predicate.

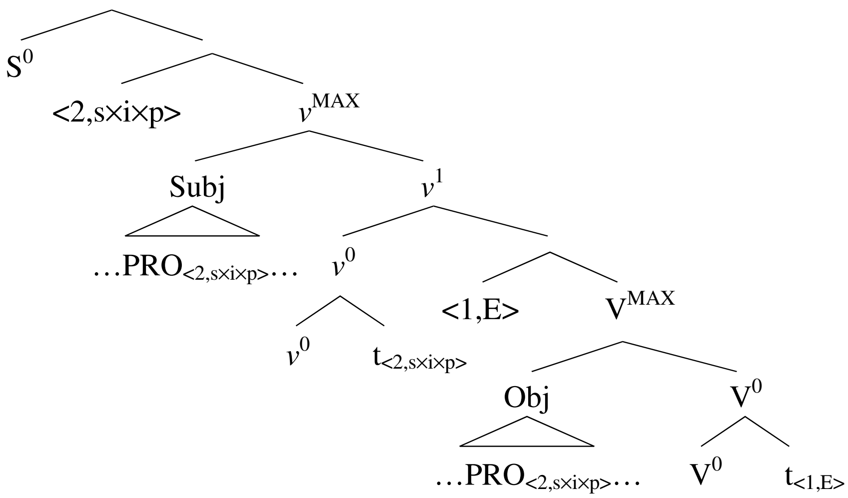
(65)



However, as I discussed in Section 2.1, this configuration is problematic, since the tense morphology is not right next to the head of the main predicate, which would be a natural condition for morphological merger to take place. Thus, the tense morphology should not be outside the VP, and even if the tense morphology is inside the VP, a noun phrase should not intervene between the head of the predicate and the tense morphology. In my framework, the time argument of the main predicate is represented by the trace that is introduced by the head movement of T. Recall that the sister of the trace is one of the heads that have comprised the head complex for this clause. This is S^0 in the derivation I proposed in Section 3.2, but if one assumes that S does not exist, then the sister of the trace will be v^0. Whichever it may be, in my semantics model, this head, which takes the trace as its first argument, relates the time denoted by the trace to the event described in the clause. Therefore, the time argument of the main predicate gets saturated the moment it is introduced, and so a noun phrase that is an argument of the main predicate is never be able to combine with the denotation of the main predicate whose time argument is unsaturated. Thus, Kusumoto's approach is not compatible with my model either syntactically or semantically.

We want to account not only for the dependence of the temporal interpretation of cardinal noun phrases, but also for the dependence of the world and spatial interpretations of cardinal noun phrases. To this end, I propose that cardinal noun phrases have a syntactically present situation argument[2] (I represent it as a PRO), which is co-indexed with the situation argument of the main predicate and so is obligatorily bound by the situation binder in the main clause at LF. (66) illustrates an LF where both subject and object noun phrases are cardinal. Since the situation binder created by the head movement of S^0 in the main clause is <2,s×i×p>, each of the noun phrases contains a variable indexed as <2,s×i×p>, and these variables and the situation argument for the main predicate (i.e., the trace of the head movement of S^0) are simultaneously bound by the situation binder.

(66)



To account for the fact that the situation variable in cardinal noun phrases is obligatorily bound by the situation binder introduced by the head movement of S^0 in the main clause, I propose that:

(67) The situation variable inside a cardinal noun phrase is an anaphor.

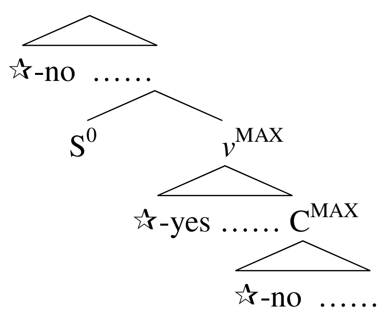
2. Kusumoto claims that noun phrase do not have a syntactically present time argument, since this assumption can explain the lack of certain readings in temporal donkey sentences and *ing*-participle clauses that would otherwise be expected. In the hope that the lack of these readings will be ruled out by some other independent reasons, I leave this issue for future research.

Anaphrs are represented by reflexives and A-traces, and they follow Condition A of Binding Theory, which states that anaphors must be locally A-bound (thus, perhaps the situation variable inside cardinal noun phrases should not be a PRO, but I continue to take it to be a PRO for convenience). Since heads are not generally considered to be A positions, binding by the S^0 head may not count as A binding. Then, (67) should be understood by analogy with A binding of nominal anaphor. Of course, if we admit that heads (or at least some heads including S^0) are A positions as well, the situation variable inside cardinal noun phrases will be an anaphor in its strict sense. Either way, the essential idea is that the situation variable inside a cardinal noun phrase has to be bound in the clause in which the noun phrase has originated, and so cannot be bound long distance by some S^0 in a higher clause. This forces the world, time and space interpretations of a cardinal noun phrase to be dependent on the corresponding interpretations of the main predicate. One might find it strange that in (66), movement is licensing new bindings that did not obtain before the movement. Although such a configuration would result in crossover effects with A-bar movement, A movement allows such a configuration as the following example shows:

(68) John $\lambda 1$ seems to himself$_1$ [t_1 to be sick].

Here, the A movement of *John* to the matrix SpecTP licenses A binding of the anaphor *himself*.

This account of cardinal noun phrases proposed above is in line with the scope approach that Musan (1995) considers and rejects (see the beginning of this section and Appendix). Since the situation varaible inside cardinal noun phrases is an anaphor, it must be bound by a local binder introduced by head movement of S^0. Let us focus on the *v*P whose sister is the S^0 that has introduced the binder that binds the situation varaible inside a cardinal noun phrase, and call it the cardinal noun phrase's interpretation *v*P. If a cardinal noun phrase moves out of its interpretation *v*P, then, since the *v*P's sister is the S^0 that has introduced the binder, the situation variable in the noun phrase will end up above this S^0, and so cannot be bound. Also, when a cardinal noun phrase's interpretation *v*P has a subordinate CP, the cardinal noun phrase should not be inside the CP, since then, the binding would cross a clausal boundary and therefore would not be local. (69) illustrates where a cardinal noun phrase can be at LF. Here, the position of the cardinal noun phrase is shown by ☆ with yes or no indicating whether or not that position is legitimate, and the v^{MAX} is the cardinal noun phrase's interpretation *v*P.

(69)


Thus, the account is compatible with Diesing's Mapping Hypothesis, according to which, cardinal noun phrases should be in *v*P and not outside *v*P (e.g., at SpecTP) at LF. Note that such an LF positional restriction is shared by nominal anaphors. In (70a), the anaphor *himself* is outside *v*P and so it cannot be A bound, and (70b) shows that *himself* cannot be A bound across a clausal boundary:

(70) a. *Himself likes John.

b. *John thinks that himself is intelligent.

Now, recall that when we discussed the Thesis of World Dependence in the previous section, we saw that a cardinal noun phrase has to receive an opaque reading when embedded under an intensional operator. This means that a cardinal noun phrase could not be raised out of the *v*P in which it has originated and end up inside the *v*P of a higher clause, since then, it would receive a transparent reading. In other words, a cardinal noun phrase's interpretation *v*P should always be the *v*P in which it has originated. However, if it should be maintained that the situation variable inside cardinal noun phrases is an anaphor in the same way as nominal anaphors are, this view would be challenged. We have ECM constructions, in which the subject of an embedded clause can be A bound by an element in the matrix clause. An example is given below:

(71) John considers himself to be intelligent.

Here, *himself* is bound by *John*, seemingly across a clausal boundary. If the biding of the situation variable is A binding involving the same binding domain as A binding of nominal anaphors, a cardinal noun phrase that is the subject of the embedded clause of an ECM verb should be able to receive a transparent and yet cardinal reading. Thus, if *a student* in the sentences in (72) has a transparent and cardinal reading, then this prediction will be borne out.

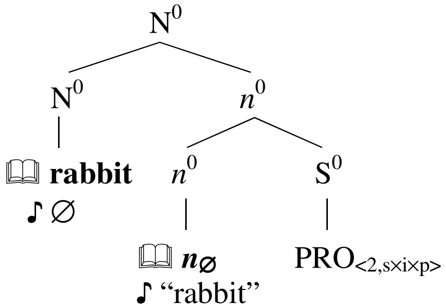
(72) a. John considers a student to be responsible for the failure.

b. John wants a student to come to the party.

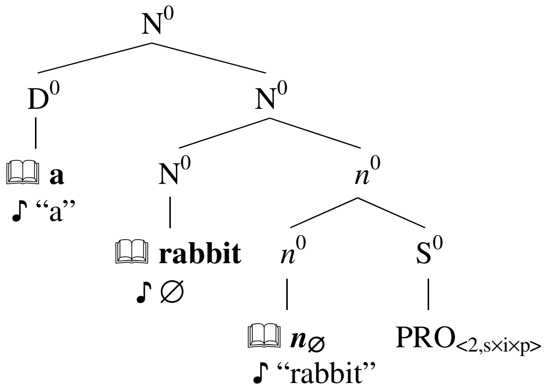
My intuitions are not clear on this, but it seems that the types of available readings do not vary depending on whether the verb is an ECM verb or takes a CP complement. For now, I leave this as an open question, and I want to investigate this point in more detail in future research.

Now, let us see what the internal structure of cardinal noun phrases is like. It is natural to assume that the situation variable in cardinal noun phrases is the PRO (or unquantifying) version of S^0, because the binder introduced by head movement of S^0 abstracts over situations. Then, the derivation of cardinal noun phrases will be similar to that of presuppositional noun phrases, except that there are less heads involved.

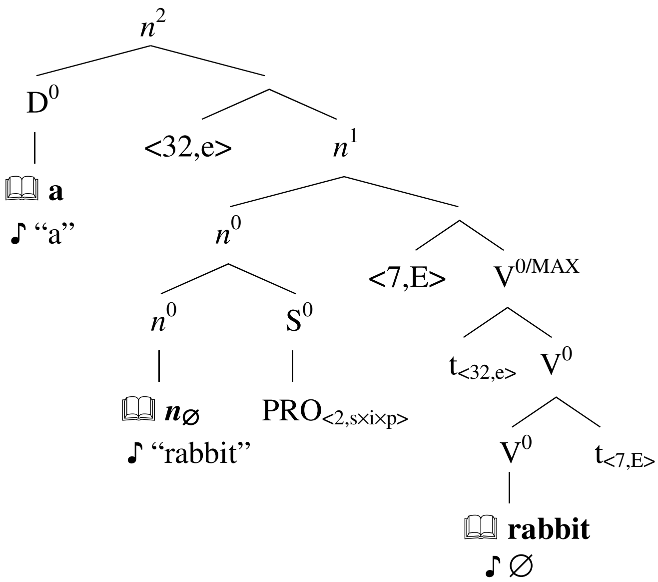
Let us take the derivation of *a rabbit* (when it is cardinal) for instance to illustrate how the derivation of a cardinal noun phrase proceeds. First, the head complex in (73) is formed:

(73) 

Since *rabbit* is an "unaccusative" noun, the determiner *a* is merged in the "object" position (i.e., inside merged with N^0):

(74) 

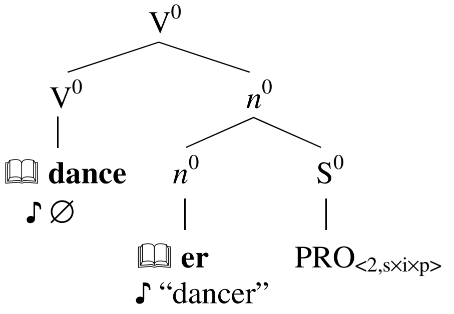
This is followed by head movement of n^0 and movement of the determiner to a Spec*n*P position:

(75) 

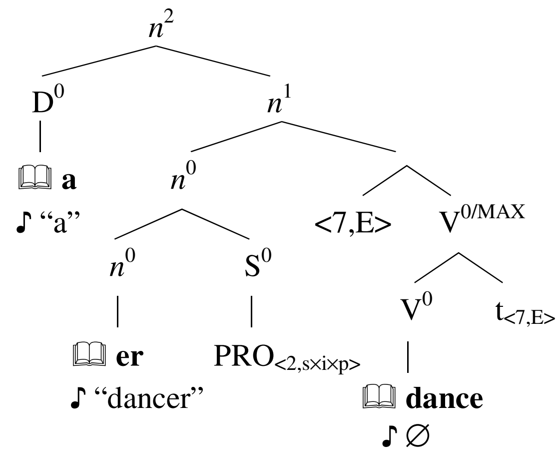
The movement of the determiner is necessary for the proper semantic interpretation and it could be argued to be a case of A movement.

Let us now consider the derivation of *a dancer* (when it is cardinal) to see the case of a cardinal noun phrase whose determiner receives the agent role. The derivation starts with the head complex in (76a), and after the head movement of n^0, the determiner is merged as a Spec*n*P as in (76b):

(76) a.



b.



Note that unlike presuppositional noun phrases which TPs, cardinal noun phrases are *n*Ps. I also refer to these *n*Ps as noun phrases.

Now, if we give the lexical entry in (77a) to the cardinal determiner *a*, the denotation of *a rabbit* with the structure in (75) is calculated as in (77b):

(77) a. $[\![\mathbf{a}]\!]^{c,a} = \lambda f\in D_{et}.\ [\lambda g\in D_{et}.\ \exists x\ [f(x) \wedge g(x)]]$

b. $[\![\textbf{a rabbit}]\!]^{c,a}$
$= \lambda g\in D_{et}.\ \exists x\in D_e\ [\exists e\in D_E\ [M(e) \supseteq a(<2,s\times i\times p>) \wedge \text{Theme}(e)(x)$
$\wedge\ \text{RABBIT}(e)]\ \wedge g(x)]$

Thus, unlike the denotation of a presuppositional noun phrase, the denotation of a cardinal noun phrase does not contain quantifications over times and spaces.

3.4.3 Some remaining problems

To end Section 3.4, I would like to point out two problems with my account.

One problem concerns why presuppositional noun phrase should be outside the *v*P at

LF. We have seen above why cardinal noun phrases have to stay inside the *v*P. Although this is compatible with Diesing's Mapping Hypothesis, Diesing's Mapping Hypothesis also says that presuppositional noun phrase have to be outside the *v*P at LF. Apparently, my model offers no account for this. Since presuppositional noun phrases come with their own quantifications that bind their temporal and space arguments and the world variable is the only argument that needs binding from outside, they should be able to appear anywhere in the tree, as long as the world variable gets bound properly. For subject noun phrases, perhaps we can hypothesize that they moves to SpecTP (presumably for case reasons) and remain there unless there is a reason to reconstruct them back into the *v*P as in the case of cardinal noun phrases. However, if object presuppositional noun phrases should also be outside *v*P, the above kind of hypothesis would not be sufficient unless we also assume that there is a syntactic reason to move object noun phrases outside *v*P such as checking some feature.

The other problem concerns why cardinal noun phrases have to come with a situation anaphor. Right now, it is a mere stipulation. One possible theory would be to say that the situation anaphor inside cardinal noun phrases is actually a trace of the moved S^0 in the main clause. That is, when the S^0 that is to undergo head movement is merged with v^0, it is at the same time merged with the n^0 in the cardinal noun phrase as well. This would require a single node to be able to merge with multiple nodes and so would complicate the syntax significantly, however.

A related question is whether or not the world varialbe in presuppositional noun phrases is an anaphor. Consider (78):

(78) Mary hopes that a friend of mine will win the race.

Note that *a friend of mind* is presuppositional here. Heim's class notes (1999) (which she has built on Kai von Fintel's notes which in turn incorporated noted by Heim and Angelika Kratzer) points out that in addition to the *de dicto* reading in (79a) and the specific *de re* reading in (79b), (78) has a non-specific *de re* reading, which is described in (79c):

(79) a. For every world w that is compatible with what Mary hopes in the actual world w*, there is an individual x such that x is a friend of mine in w and x will win the race in w.

b. There is an individual x such that x is a friend of mine in the actual world w* such that for every world w that is compatible with what Mary hopes in w*, x will win the race in w.

c. For every world w that is compatible with what Mary hopes in the actual world w*, there is an individual x such that x is a friend of mine in w* and x will win the race in w.

Heim explains that in order to account for this non-specific *de re* reading, unless we posit movement of NP to the exclusion of the determiner (i.e., movement of *friend of mine* only) to scope it out of the embedded clause, we need a long-distance binding of the world argument in the noun phrase *a friend of mine*. If such long-distance binding is necessary, then the world variable I posit in presuppositional noun phrases cannot be an anaphor.

Then, a picture emerges in which the world variable in presuppositional noun phrases is not an anaphor unlike the situation variable in cardinal noun phrases. How could this be derived? Recall our discussion in the previous subsection that (A) binding seems to be possible only when the bound variable is in *v*P save cases of ECM constructions. If this is a correct generalization, then it would automatically follow that the world variable may not be an anaphor, as it is a C^0 and thus is outside the *v*P in the internal structure of a presuppositional noun phrase (see subsection 3.4.1).

Thus, we have seen that my model give rise to a number of questions to which we have no conclusive answers. I hope to investigate these deeper in future research.

3.5. A scope puzzle with cardinal determiners

In this section, I show that cardinal noun phrases whose determiner is something other than *a* present a scope puzzle.

To begin with, consider the following sentence, where *three bachelors* is cardinal:

(80) Mary kissed three bachelors yesterday.

As discussed in the previous section, cardinal noun phrases remain inside *v*P so that their situation argument will be bound by the binder introduced by the head movement of S^0. Assuming that $PRO_{<13,i>}$ is at SpecTP and $PRO_{<11,p>}$ is at SpecSP, the truth condition of (80) is computed to be in (81) under indexical assignment *d* and variable assignment a:

(81) $\forall w \in D_s\ [@(w) \rightarrow \exists t \in D_i\ [t < d(\text{NOW}) \wedge t \subseteq a(<13,i>) \wedge \exists s \in D_p\ [s \subseteq a(<11,p>)$
$\wedge \exists x \in D_e\ [\exists e' \in D_E\ [M(e') \supseteq (w,t,s) \wedge \text{Theme}(e')(x) \wedge \text{Bachelor}(e')]$
$\wedge \exists e \in D_E\ [e \text{ occurs in } (w,t,s) \wedge \text{Agent}(e)(\text{Mary}) \wedge \text{Theme}(e)(x) \wedge \text{KISS}(e)]]]]$

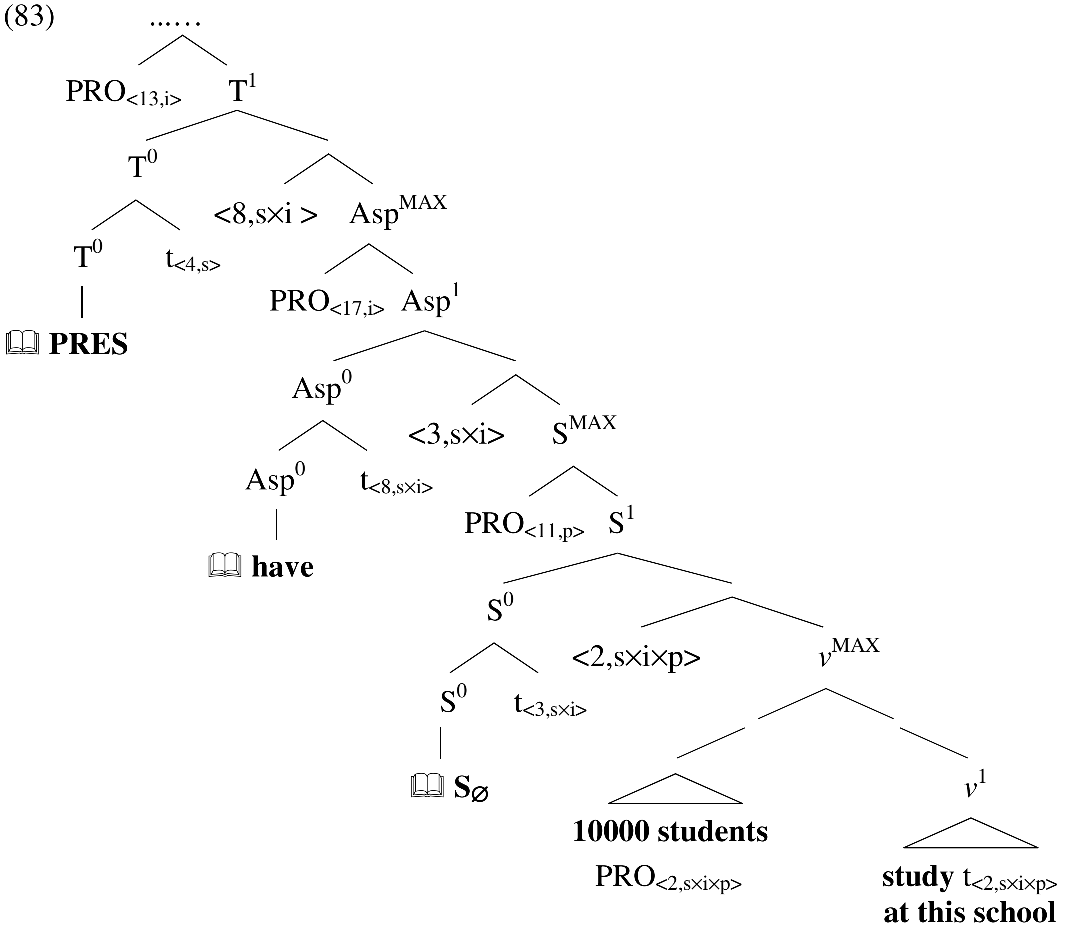
Thus, (80) would involve three individuals that are bachelors in the same situation as Mary's kissing took place in. However, Mary might well have kissed these three guys separately in different places at different times of yesterday, and such a case seems to conflict the truth condition in (81). One might think that this conflict may be resolved by thinking of a big situation that contains the three different kissing situations in which she kissed each of the bachelors. This does not work, however. Imagine that each bachelor got married right after Mary kissed him, and that when she kissed the next one, the previous one was no longer a bachelor. In this scenario, if we think of a big situation containing all the three kissing situations, at least two of the guys would fail to be a bachelor in this big situation, because according to my semantics, being a bachelor in a situation requires being a bachelor throughout the situation.

(82) makes the same point perhaps more clearly.

(82) 10000 students have studied at this school.

Assuming that *have* heads an AspP which lies between TP and SP, we would have the following LF for (83):

(83)



Here, I am assuming the following lexical entry for *have*:

(84) $[\![\textbf{have}]\!]^{c,a} = \lambda\tau\in D_{s\times i}.\ [\lambda f\in D_{<s\times i,t>}.\ [\lambda t' \in D_i.$
$\exists t\in D_i\ [t < \text{Time}(\tau) \wedge t \subseteq t' \wedge f((\text{World}(\tau),t))]]]$

Time(τ) and World(τ) are functions such that for any $\tau = (w,t)$, where $w\in D_s$ and $t\in D_i$, Time(τ) = t and World(τ) = w. Thus, *have* basically just shifts the time given by the temporal quantifier of T^0 towards the past. Given the tree and the semantics of *have*, (82) would then roughly mean that there is some situation s in the past such that for 10000 individuals x that are students in s, there is an event of x studying at this school that occurs in s. Then, one would need to think of a situation in which all these 10000 individuals are student simultaneously. However, people keep on entering school and graduating, and so naturally, there would be people who have studied at this school but have never been students at the same time. Thus, the truth condition we obtains does not seem appropriate.

The above cases both involve a determiner that is a number bigger than one such as *three* and *10000*, and the quantified individuals need to be evaluated with respect to different situations. In (80), the kissed individuals are bachelors in different situations, and in (82), people who studied at the school are students in different situations. To achieve this, it is required that the quantified noun phrase take scope over the quantifier over situations, which is achieved by the combination of T^0 (and Asp^0) and S^0. Thus, the quantified noun phrase must c-command the T^0 that has moved by head movement. However, since the quantified noun phrase is cardinal in these cases, their situation variable must be bound by the situation quantifier in the main clause. In other words, the noun phrase must be c-commanded by the S^0 that has moved by head movement. Since the T^0 c-commands the S^0, this appears to be a scope paradox.

Now, consider the following:

(85) Mary kissed no bachelor yesterday.

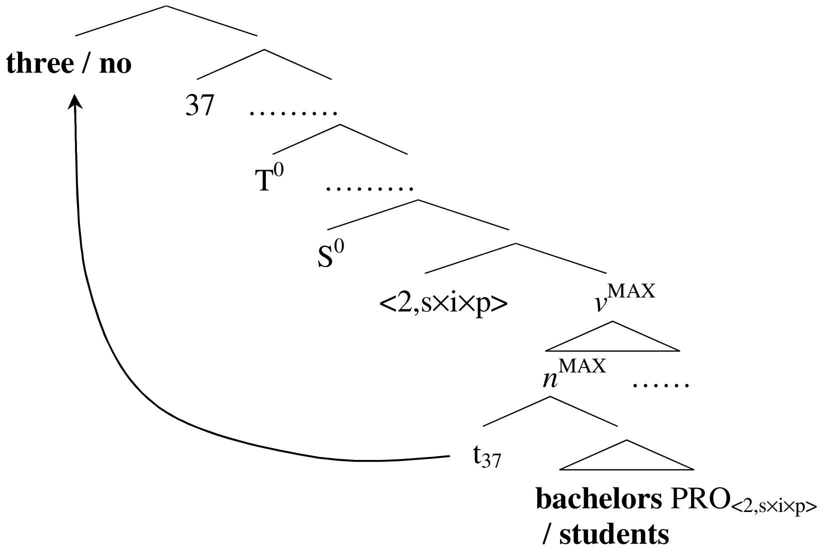
Here, since *no bachelor* is cardinal, it remains inside *v*P. Then, according to our semantics, (85) would be true if there is some situation in the past in which Mary kisses no bachelor. Similarly, many sentences such as (85) that involve a cardinal noun phrase with *no* should be trivially true, but that contradicts our intuition. The appropriate interpretation of (85) should be something like the following:

(86) There is no individual x such that there is a situation s in yesterday such that x is a bachelor in s and Mary kissed x in s.

Thus, just like the cases with a determiner that is a number bigger than one, it seems that the quantified noun phrase needs to take scope over the situation quantifier, but again, since the noun phrase is cardinal, it would lead to a scope paradox for the same reason.

Thus, once we look at cardinal noun phrases whose determiner is other than *a*, we see that there appears to be a scope paradox. An obvious proposal to solve this is scoping out only the determiner of the cardinal noun phrase to a SpecTP position, leaving the residue of the noun phrase behind. On this approach, the LF of the sentences that we have looked at above would look roughly like (87):

(87)



However, merely moving out the determiner does not work. In the previous section, I proposed the following lexical entry for the cardinal determiner *a*:

(88) $[\![\mathbf{a}]\!]^{c,a} = \lambda f\in D_{et}. [\lambda g\in D_{et}. \exists x\ [f(x) \wedge g(x)]]$

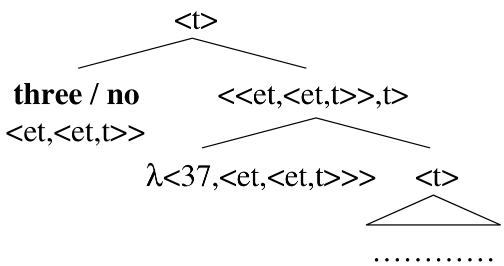
Thus, the denotation of *a* is of type <et,<et,t>>. If this is the semantic type of cardinal determiners in general, the denotations of *three* and *no* would be as follows:

(89) a. $[\![\mathbf{three}]\!]^{c,a} = \lambda f\in D_{<et>}. [\lambda g\in D_{<et>}. 3x\ [f(x) \wedge g(x)]]$

b. $[\![\mathbf{no}]\!]^{c,a} = \lambda f\in D_{<et>}. [\lambda g\in D_{<et>}. \neg\exists x\ [f(x) \wedge g(x)]]$

Since the top node of the tree in (87) should be of type t, the binder 37 should then be abstracting over elements of type <et,<et,t>>, as (90) below shows:

(90)



Therefore, the trace of determiner (i.e., t_{37}) should be of type <et,<et,t>>. Then, since it is the case that the determiner is leaving a trace that is of the same type as itself, this movement has no semantic consequence, and thus the intended interpretation will not be obtained.

We are thus led to conclude that we have to revise the denotations in (89). Specifically, we need to extract only the quantificational part of the denotation as the proper denotation of the moving element. Thus, the denotations of *three* and *no* should be as follows:

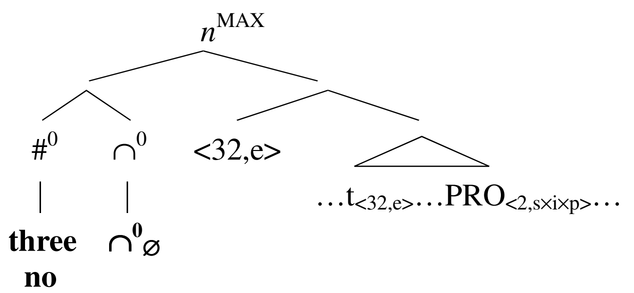
(91) a. $[\![\textbf{three}]\!]^{c,a} = \lambda f \in D_{<et>}.\ 3x\ [f(x)]$

b. $[\![\textbf{no}]\!]^{c,a} = \lambda f \in D_{<et>}.\ \neg\exists x\ [f(x)]$

three and *no* are quantifiers of type <et,t>, and they leave a trace of type e behind when they move. The difference between the old denotations in (89) and the new denotations in (91) consists in that unlike the old ones, the new denotations do not take two properties of individuals and intersect them. This actually makes sense, given that we are considering cardinal determiners. The denotations in (89) are appropriate for presuppositional determiners. Since presuppositional noun phrases have a restrictive clause (Heim (1982), Diesing (1992)), the denotation of a presuppositional determiner needs to take two properties of individuals as its arguments, of which one functions as the resource domain. On the other hand, cardinal noun phrases do not have a restrictive clause, and therefore the denotation of a cardinal determiner must have only one argument that is a property of individuals.

How can we implement this idea in syntax, then? One hypothesis would be to propose that what we originally thought of as a determiner is actually the combination of the numerical part such as *three* or *no*, whose category I tentatively call #, and a new item whose category I tentatively call $\cap$. On this approach, the internal structure of a cardinal noun phrase would look something like the following:

(92)



At LF, $\#^0$ moves out and leaves a trace of type e. The denotation of $\cap^0_{\varnothing}$ is given as follows:

(93) $[\![\cap^{\mathbf{0}}]\!]^{c,a} = \lambda x \in D_e.\ [\lambda f \in D_{<et>}.\ [\lambda g \in D_{<et>}.\ f(x) \wedge g(x)]]$

While this hypothesis works semantically, it is syntactically problematic. We are assuming that only *three* or *no* moves to SpecTP in the main clause, leaving $\cap^0_{\varnothing}$ and everything else in n^{MAX} behind. Since the n^{MAX} is the specifier of *v*P or VP, this is a movement out of a specifier. However, we do not seem to know any movement of this kind. Phrasal movement of such kind is supposed to be impossible (Huang (1982)), and as I discuss extensively in Section 4, head movement is not possible out of a specifier. However, I would like to refer to Diesing (1992), who discusses the extraction out of cardinal noun phrases is possible while extraction out of presuppositional noun phrases is impossible. Recall that in my derivation model, both the subject and the object are specifiers. Thus, as long as it is possible to extract out of a cardinal object noun phrase, we will have to admit the existence of extraction out of a specifier. If the generalization of extraction that Diesing discusses applied to $\#^0$ as well, then the movement of $\#^0$ would be expected to be possible. Of course, this does not solve all the problems. In fact, if this line of approach is correct, there should be no reason that $\#^0$ can move out of a cardinal subject noun phrase, since extraction out of a subject is known to be impossible (Huang (1982)), and moreover, my model does have an account of why extraction out of a subject is impossible.

Thus, as things stands now, a lot of problems remain regarding the treatment of cardinal noun phrases with a determiner other than *a*.

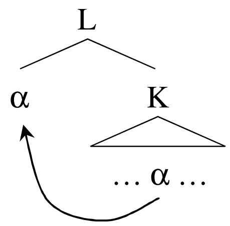
Chapter 4 Head Movement and Phrase Structure

In this chapter, I try to pin down characteristics of the syntax that is need by my model of head movement. I define a new Phrase Structure, explore the appropriate locality for head movement and propose an axiom that only maximal projections may move. I show that the observed properties of head movement are derived from these. At the end of the chapter, we finally return to Percus's problem and see that it is a necessary consequence of our syntax.

4.1 Head movement and locality

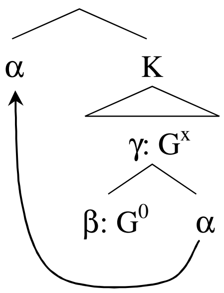
One important characteristic of head movement is that the moved element is projected. (1) below illustrates the general configuration of a movement.

(1)



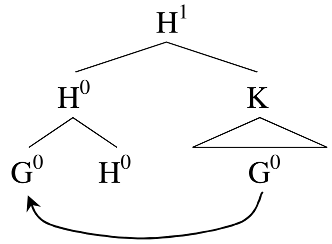
In the case of a phrasal movement, the moved element α is a phrase, and L is a projection of K. Thus, α has moved to a specifier. On the other hand, in the case of the head movement as proposed in my model, α is a head and L is a projection of the moved element α, and not of K. This is a configuration that Chomsky (1994, 1995) claims does not exist. In the conventional model of head movement, α should be the head of K, since Head Movement Constraint requires that head movement be strictly local. Then, (1) should be a case of self adjunction, and Chomsky argues that it should not be permitted. In my model, by contrast, assuming that α is a head and that it is projected after movement does not necessarily yield a case of self adjunction. In my model, α may be merged with another head β to form a projection of β. Let β be G^0 and its mother be $\gamma = G^x$ (where x = 0 or 1; we discuss this shortly). Then, K could be a projection of β, but not of α, and so the configuration is not a case of self adjunction as shown below:

(2)



The configuration in (2) requires us to revise the locality condition on head movement, however, because in the conventional theory, only the head of K should be able to undergo head movement and that is not α in (2). Nonetheless, we still want to capture intuitively the same locality condition in our model as in the conventional one. Let's now look at what corresponds to (2) in the conventional model, where H^0 corresponds to α in (2):

(3)



Here, the locality condition requires that G^0 be the head of K, i.e., $K = G^n$. What moves in (3) is G^0, but in (2), what moves is α, the sister of G^0. Nevertheless, the movements in (3) and (2) correspond to each other. Therefore, if the locality condition says that only G^0 is eligible for head movement in (3), it should be translated as this in my model: only the sister of G^0 is eligible for head movement in (2). Since $\beta = G^0$ is the head of K (provided that β is the only head of K; see below), the locality condition we want is that the moved head be the sister of the head of K.

That α be the sister of the head of K is only a necessary condition, and not a sufficient one. Suppose that x = 1, i.e., $\gamma = G^1$ in (2). Then, α would be a complement or a specifier of β. If α were local to K in this case, it should be able to undergo head movement, but complements and specifiers are not known to undergo head movement. Chomsky (1994, 1995) discusses that clitics may be analyzed as heads and maximal projections at the same time, but they never get projected after movement. If we allowed specifiers to undergo head movement, clitics would be able to undergo head movement and be projected after movement, contrary to the fact. This consideration leads us to conclude that x = 0, i.e., γ is G^0, another head. Note that since specifiers and complements are not local and cannot undergo head movement, no more is head movement out of a specifier or a complement, which is not attested, possible. This is the reason I assumed that merger of two heads yielded another head in Section 2.2. In sum, we also need to add the condition that the mother of the moved head α be itself a head. However, since both β and γ could be considered as heads of K, α is no longer the sister of <u>the</u> head of K. Thus, the necessary and sufficient locality condition would be something like the following:

(4) Locality in head movement (to be revised in Section 4.4)
Given K, if there are nodes α, β and γ such that α and β are the daughters of γ, and β and γ are both heads of K, then only α is local to K with respect to head movement.

Although this captures exactly what undergoes head movement, it looks like the definition of locality is getting rather complex and stipulative. In section 4.4, I replace this with a simpler locality condition and show that (4) is derived. Given the locality, head movement can now be defined as follows:

(5) Definition of head movement
Given K, let α be a head that is local to K. Then, α may be internally merged with K.

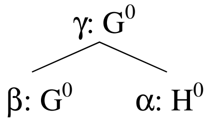
This definition makes use of the word "head", which is defined in Section 4.3.

4.2 Semantic types and syntax

4.2.1 Is the complex head a projection?

Consider a complex head $\gamma = G^0$ whose daughters are $\alpha = H^0$ and $\beta = G^0$:

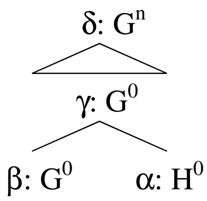
(6)



Here, a question arises as to the syntactic status of γ. Is γ a projection of β? Or is it the case that α is an adjunct and so β and γ are segments? In the conventional model of head movement, head complexes are formed by adjunction as a result of head movement. Thus, γ would be formed by adjunction of α, which has moved by head movement. Therefore, γ should not be a projection of β in the conventional thoery. I argue, however, that in my model, γ is a projection of β.

To see this, let us examine the semantic types involved in the complex head formation. Suppose that α in (7) undergoes head movement and is merged with δ. Since α is local to δ, $\delta = G^n$ (i.e., a projection of γ).

(7)



Here, it might be the case that $\delta = \gamma$. According to the semantics I have developed in Chapter 3, α is either a PRO or a quantifier that undergoes head movement. Furthermore, for a given syntactic category H, if an H^0 that is a PRO denotes an element of type σ, then an H^0 that is a quantifier quantifies over elements of type σ and so leaves behind a trace of type σ, and vice versa. Then, no matter whether α in (7) is a PRO or a quantifier, β has to denote a function whose first argument is of some semantic type σ. Notice here that type σ should be neither t (i.e., truth-values) nor function types such as <e,t>, etc., since as we saw in Chapter 3, when α is a quantifier, it quantifies over events, situations, etc., and never over truth-valued or functions. Let us define simple type as follows:

(8) Definition

(i) e, E, p, i and w, are **simple types**.

(ii) For any **simple types** x and y, $x \times y$ is a **simple type**.

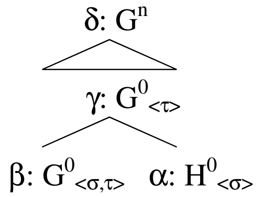
Then, we can say that type σ is a simple type. Thus,

(9) Suppose that α, β and γ are heads, where γ is the mother of α and β, and $\alpha = H^0$ and $\beta = G^0$. Suppose further that $\gamma = G^0$. Then, β denotes a function whose first argument is of a simple type.

It is not clear if this should be derived from some more fundamental principles. It may just be a coincidence due to lexical entries, but I henceforth assume that (9) is a correct generalization.

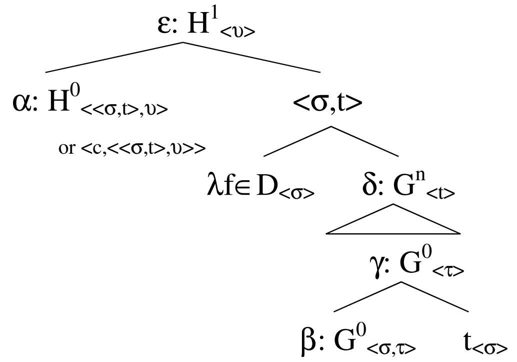
Now, let us first consider the case where α in (7) is a PRO. Assume that this PRO's denotation is of type σ, which is a simple type. The type of β should be of type $<\sigma,\tau>$ for some semantic type τ so that it acts as a function taking α as its argument as shown below:

(10)



Now, let us assume that α in (7) is quantifier. Since an H^0 that is a PRO denotes an element of type σ, an H^0 that is a quantifier quantifies over elements of type σ. Thus, α is of type $<<\sigma,t>,\upsilon>$ for some semantic type υ in cases where Functional Application is employed to interpret the mother node of the moved α, and is of type $<c,<<\sigma,t>,\upsilon>>$ in cases where Intensional Functional Application is employed. When α moves, it leaves behind a trace of type σ. β has the same denotation as in the above case, and so is of type $<\sigma,\tau>$. In my model, complements always denote a truth-value. In this case, δ becomes the complement of α after the head movement. Predicate Abstraction is employed to interpret the structure, and since α is of type $<<\sigma,t>,\upsilon>$ or $<c,<<\sigma,t>,\upsilon>>$, δ must denote a truth-value as illustrated below:

(11)

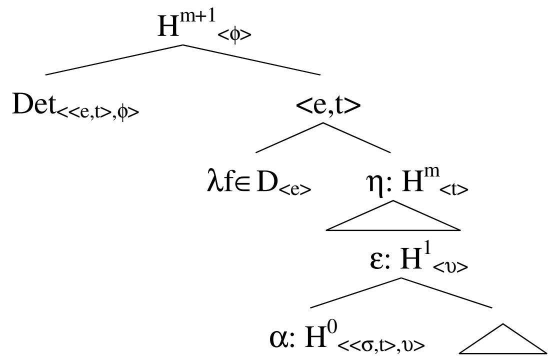


It then turns out that τ and υ belong to a specific group of semantics types. Since γ, which is of type τ, is a head of δ, which is of type t, denotations of type τ must be either truth-values (i.e., δ = γ) or functions that eventually produce a truth-value. And actually, the same generalization holds for υ as well. In other words, if the root node ε does not denote a truth value, it must denote a function that eventually produces a truth-value.

This is obvious when α dominates a head that undergoes head movement, since then, exactly the same argument can be made for ε as for γ.

Now, assume that α does not dominate a head that undergoes head movement. If the structure in (8) is the main clause, the generalization must hold since sentences denote truth-values. Now, suppose that the structure in (8) is part of a noun phrase. Since no head movement takes place out of α, α will have a determiner as a specifier which has moved from inside the complement of α. Since a determiner is of type <<e,t>,ϕ> (ϕ = <<e,t>,t> for quantifying determiners and ϕ = e for the definite article), its sister should denote a truth-value:

(12)



Thus, again, denotations of type υ must be either truth-values (i.e., η = ε) or functions that eventually produce a truth-value.

Let us now define T-type as follows:

(13) Definition

(i) t is a **T-type**.

(ii) For any **T-type** x and any semantic type y, <y,x> is also a **T-type**.

Note that a simple type is never a T-type. With this, we can generalize that τ and υ are both T-types. Thus,

(14) Suppose that α, β and γ are heads, where γ is the mother of α and β. Let $\alpha = H^0$ and $\beta = G^0$. Suppose further that $\gamma = G^0$. Then, β's denotation is of type $<\sigma,\tau>$, for some simple type σ and some T-type τ.

We can now see that it is misleading to think that α is an adjunct to β in (6). Adjuncts are dispensable constituents such as adverbials, and semantically speaking, Predicate Modification should be employed in interpreting structures containing them. In the case of head complex formation in my model in (6), Functional Application is applied to β and α or the trace of α, and so α is not dispensable since β takes α or its trace as its argument. Therefore, α should not be considered as an adjunct, and this leads us to conclude that γ is a projection of β. α is a "complement" of β in the sense that α is not omissible, but α is not a complement in the syntactic sense, since complements do not undergo head movement. More precisely, if α were a complement, γ would no longer be a head, so $\gamma = G^1$, but this is not compatible with the locality we observed in (4) above. To sum up, γ is a projection of β and is still a head.

4.2.2 What is projected?

We have established above that a complex head is a projection of one of its sisters. Suppose that $\alpha = H^0$ and $\beta = G^0$ are sisters and yield a head complex γ:

(15)
$$\begin{array}{ccc} & \gamma & \\ \diagup & & \diagdown \\ \beta: G^0 & & \alpha: H^0 \end{array}$$

γ is a projection of either β or α, i.e., either $\gamma = G^0$ or $\gamma = H^0$, but how do we know which γ is going to be? An easy way to do this is to resort to syntactic selection. It is usually assumed, for example, that verbs select for NPs/DPs, and therefore the combination of a verb and an NP/DP will be a projection of the verb. In a similar fashion, we could just assume that V selects for *v*, T selects for C, etc.

However, if we consider the semantic types involved in the configuration, we can make a straightforward generalization about what is projected without recourse to syntactic selection. In both (10) and (11), β denotes a function that takes α or its trace as its first argument, and γ is a projection of β. In addition, since we have established that a complex head is a projection of one of its daughters and since the generalization is based on semantic interpretation, it should be not only about complex heads, but about projections in general. Thus, I propose the following:

(16) If γ is a non-terminal node that is a projection of one of its daughters, then γ is a projection of the daughter that acts as the function when Functional Application is applied to them to compute γ's denotation.

I think that (16) makes sense intuitively, but in order to see if it is a tenable generalization, let us look at a similar but familiar paradigm: noun phrases and predicates of individuals.

Consider the examples below:

(17) a.

$$\gamma_{<t>}$$

α: **John**$_{<e>}$ β: **danced**$_{<et>}$

b.

$$\gamma_{<t>}$$

α: **everyone**$_{<et,t>}$ β: **danced**$_{<et>}$

In both examples in (17), β denotes a predicate of individuals (i.e, of type $<e,t>$). In (17a), α denotes an individual (type e), and in (17b), α denotes a quantifier over individuals (type $<<e,t>,t>$). In both cases, we normally take it that γ is a verb phrase and not a noun phrase. In other words, γ is β's projection. According to (16), γ is β's projection in (17a) as expected. What about (17b) then? If *everyone* is QR'd and leaves a trace of type e behind, then *danced* is going to act as a function that takes the trace of *everyone* as its argument when computing the value of γ. Therefore, γ will be β's projection in (17b). However, if *everyone* stays *in situ*, then it is going to act as a function that takes *danced* as its argument when computing the value of γ. Therefore, γ will be α's projection, but this is not a desired result.

One way to avoid concluding that γ is α's projection it to modify the generalization in (16). Recall that if γ is β's projection, then β's denotation is of type $<\sigma,\tau>$, for some simple type σ and some T-type τ. Thus, γ has at least one daughter with a denotation of that kind, but actually, it turns out that not both of γ's daughters may have such denotations. To see this, suppose to the contrary that one of γ's daughters has the denotation of type $<\sigma,\tau>$, where σ is a simple type and τ a T-type, and the other has the denotation of type $<\upsilon,\phi>$, where υ is a simple type and ϕ a T-type:

(18)

$$\gamma$$

$<\sigma,\tau>$ $<\upsilon,\phi>$

The denotation of γ should not be computed by Predicate Modification, since that would imply that one of γ's daughter would be semantically omissible and so would be an adjunct. However, as I discussed in the previous subsection, I assume that a complex head is always a projection, and so neither of γ's daughters is dispensable and Functional Application should be employed to compute γ's denotation. Now, if we try to apply Functional Application to γ's daughters in (18), we deduce that either $\sigma = <\upsilon,\phi>$ or $\upsilon = <\sigma,\tau>$. However, this contradicts the assumption that σ and υ are simple types. Hence, if γ is a head that is formed by merging two heads, exactly one of its daughters denotes a function that takes a denotation of a simple type and yields a denotation of a T-type. Then, we can now take (14) inversely to state the following:

(19) If γ is a non-terminal node that is a projection of one of its daughters, then γ is a projection of the daughter whose denotation is of type $<\sigma,\tau>$, for some simple type σ and some T-type τ.

Returning to (17b), since *danced*, and not *everyone*, denotes a function whose first argument is of a simple type, we can tell that *danced* is projected. However, it should be noted that by proposing (19), we would have to stipulate that two heads cannot yield a complex head, in cases where none or both of them denote a function whose first argument is of a simple type.

Although (19) is a correct generalization, (16) would actually suffice if we assumed that every quantifier obligatorily undergoes QR, since a moved quantifier leaves a trace that acts as a variable that is taken by the denotation of its sister as its first argument. In constructing the main clause, every head of one of the categories that comprise the main clause (i.e., V, *v*, S, T and C, hence main clause categories) that is a quantifier has to undergo QR (head movement), since they encounter a type mismatch due to the semantics I give to them. (17b) is different in that the structure is interpretable without having the quantifier QR'd. Nevertheless, even noun phrase quantifiers that could be interpreted *in situ* are obliged to move in many cases. For example, it is generally assumed that predicate-internal subjects generally undergo A movement to SpecTP. Thus, the quantifier subject in (17b) would have to move out. Also, as we saw in detail in Chapter 3, presuppositional noun phrase have to move out of *v*P and end up at SpecTP, as Diesing proposes. Thus, even objects have to move out if they are presuppositional. Then, the only case that remains to be a problem is non-subject cardinal noun phrases. It should be noted that even these can move, though not obligatorily, for reasons such as topicalization. If it turns out that non-subject cardinal noun phrases also obligatorily move out, it can be maintained that (16) suffices. In fact, as I discuss in detail in Section 4.4.2, when α moves, we immediately know that its mother is not a projection of α (see Axiom of Movement). Therefore, if every quantifier obligatorily moves out, it automatically follows that their mother (before movement) is not their projection. Thus, (16) would then be needed only to account for cases that do not involve quantifiers. It should be noticed, though, that (19) is a correct generalization, regardless of whether (16) may be sufficient or not.

4.2.3 Quantifiers and the structure of noun phrases

We have seen above that a non-terminal node that is a projection of one of its daughters is a projection of the daughter whose denotation is of type $<\sigma,\tau>$, for some simple type σ and some T-type τ. Then, what is the semantic type of the daughter that is not projected? When it is of type σ, this other daughter can be interpreted *in situ*, but when it is not of type σ and this daughter cannot be interpreted *in situ*, it moves out, leaving behind a trace of type σ.

We have looked at two such cases above. One is regarding the head complex and the other is regarding quantifying noun phrases. In the head complex, the v^0, T^0, etc. that are not projected move out because of a semantic type mismatch. If we recall their semantics we gave to them in Chapter 3, we can see that a moving head has a denotation of type

<<σ,t>,υ> for some semantic type υ, when its sister, which is projected, has a denotation of type <σ,τ>, where σ is a simple type and τ a T-type. In the case of a predicate with a quantifying noun phrase such as *everyone*, the quantifying noun phrase's sister, which gets projected, has a denotation of type <e,t>, whereas the quantifying noun phrase has a denotation of type <<e,t>,t>. Thus, the generalization seems to be that when the projected node has a denotation of type <σ,τ> (and when its sister is not of type σ,) its sister has a denotation of type <<σ,t>,υ> for some semantic type υ.

Now, let us further confirm that this generalization holds in the internal structure of noun phrases as well. In Chapter 3, I proposed that determiners are generated either inside N^{MAX} / V^{MAX} or inside n^{MAX}, depending on the noun, and that they move to a specifier position in the highest projection of the noun phrase (i.e., SpecTP in presuppositional noun phrases and Spec*n*P in cardinal noun phrases) except in case the determiner is generated at Spec*n*P of a cardinal noun phrase (movement is not necessary in this case). As we saw in Chapter 3, what the determiner is merged with is either N^0/V^0 or v^1, and this N^0/V^0 or v^1 gets projected. This N^0/V^0 or v^1 denotes properties of individuals i.e., it has a denotation of type <e,t>. This conforms to (19). Furthermore, the determiner's denotation is either of type <<e,t>,<<e,t>,t>> (for determiners like *every*) or <<e,t>,e> (for the definite article *the*). In either case, the determiner has a denotation of type <<e,t>,υ> for some semantic type υ. This is what we have expected.

I'd like to refer to nodes whose denotation is of type <<σ,t>,υ> as quantifiers over element of type σ. Then, quantifiers over elements of type σ should leave behind a trace of type σ when they move. Thus:

(20) Definition
Q is a quantifier if and only if Q's denotation is of type <<σ,t>,τ> for some simple type σ and some semantic type τ. When Q moves, it leaves behind a trace of type σ.

To sum up, a node whose denotation is of type <σ,τ> for some simple type σ is merged with either a node whose denotation is type σ or a node that is a quantifier over elements over type σ in the sense of (20).

4.3 Phrase Structure

We have seen in the previous section that when two heads are merged, they may yield another head which is a projection of one of its sisters. Now, if this complex head formation takes place every time two heads are merged, how could we even form a phrase? By definition, all terminal nodes are heads. Then, if merger of two heads always yields another head, every node in a tree, including the whole tree itself, will be a head.

We have movement, however. Recall that in my model, the complement of a head H is a node from within which H has moved by head movement and with which H has been merged. As I briefly discussed in Section 2.2, if that is the only instance of the complement, the arbitrary distinction between the complement and the specifier in the conventional minimalist theory will disappear. This leads to a new phrase structure theory.

To begin with, let us call all X^ns phrases for n > 0, as opposed to heads, which are X^0s:

(21) a. Definition of head
α is a **head** if and only if $\alpha = X^0$.

b. Definition of phrase
α is a **phrase** if and only if $\alpha = X^n$, where $n > 0$.

Then, the complement is defined as follows:

(22) Definition
C is the **complement** of a head H if and only if H has moved from within C by head movement and been merged with C.

We know that a head and its complement constitute a phrase (i.e., X^1). Thus, head movement automatically yields a phrase. Recall that if a head and its complement constituted another head, the complement should be able to undergo head movement according to (4), which is not the case.

In addition, so-called phrasal movement (A and A-bar movement) must yield phrases, too. Suppose that β undergoes A or A-bar movement and is merged with α. Since it is A or A-bar movement, β must have extracted from inside a complement. Thus, α has a complement in it. Let $H(\alpha)$ be a head and $C(\alpha)$ the complement of $H(\alpha)$ such that $\alpha = [H(\alpha)\ C(\alpha)]$. Then, Y must have moved from inside $C(\alpha)$. By definition, $H(\alpha)$ has moved from inside $C(\alpha)$ by head movement, and therefore $H(\alpha)$ is a phrase ($= X^1$). Thus, in order for β to undergo A or A-bar movement and be merged with α, α must already be a phrase. Now, Internal Merge of β and α produces their mother, which we call γ. Since β is undergoing A or A-bar movement, we know that β becomes a specifier and so γ is a projection of α. Then, $\gamma \neq X^0$, since $\alpha = X^1$ and γ is a projection of α, unless a projection of X^n is allowed to be X^m for some $m < n$, which is not a usual assumption. Moreover, if $\gamma = X^0$, β should be able to undergo head movement according to (4), but we know that specifiers do not undergo head movement. Given the above consideration, specifiers can then be defined as follows:

(23) Definition
S is a **specifier** of a head H if and if both S's sister and S's mother are phrasal projections of H.

Let us turn to External Merge.

In complex head formation, merger of two heads by External Merge yields another head that is a projection of one of the two heads. If we in addition want to allow merger of two heads by External Merge to yields a phrasal projection (i.e., X^1), we will need some principle to predict when a complex head should be formed and when a phrase should be formed. Allowing such an ambiguity in phrase structure would be a complication, however. It would be simpler if merger of two heads by External Merge only yielded complex heads. If we take this line of approach, we can generalize that every phrase is created only by virtue of movement. More precisely, a phrase must be either formed by Internal Merge or dominate a node that is formed by Internal Merge.

Now, let us consider cases of External Merge of two phrases. Suppose that two phrases X^n and Y^m are merged by External Merge. Then, obviously, the newly formed node will be X^{n+1} when X^n gets projected, and Y^{m+1} when Y^m gets projected.

Finally, let us consider what happens when a head and a phrase are merged by External Merge. Suppose that a phrase X^n and a head Y^0 are merged by External Merge. When Y^0 is projected, it is natural to assume that the resulting node is Y^1, rather than Y^0. A question arises when X^n is projected. Is the resulting node going to be X^n or X^{n+1}? Either way, the resulting node is a phrase and it is hard to see what differences this choice might bring about. If clitics are heads as Chomsky suggests, then, heads should be able to merge onto a phrase in cases where a clitic is cliticized onto a phrase. In such cases, it seems more natural to assume that the cliticization does not change the projection level of the phrase onto which it gets cliticized. This would enable us to make the generalization that when an externally merged head is not projected, the resulting node has the same projection level as the sister of the externally merged head (recall that when two heads are externally merged in the head complex formation, the resulting node is a head). Therefore, although we do not have any strong evidence to support it, I assume that a phrase X^n and a head Y^0 are merged by External Merge and X^n is projected, the resulting node is X^n.

Then, a new phrase structure can be formulated as follows:

(24) Phrase Structure

External Merge

Clause I: complex head formation and some other cases

Let $\alpha = X^n$ and $\beta = Y^0$, where $n \geq 0$. If α and β are merged by External Merge, and α is projected, the resulting syntactic object is X^n, a projection of α.

Clause II: phrase formation

Let $\alpha = X^n$ and $\beta = Y^m$, where $n \geq 0$ and $m > 0$. If α and β are merged by External Merge and α is projected, the resulting syntactic object is X^{n+1}, a projection of α.

Internal Merge

Clause III: head movement

Let $\alpha = X^n$ and $\beta = Y^0$, where α dominates β. If β undergoes head movement and is merged with α, the resulting syntactic object is Y^1, a projection of β.

Clause IV: specifier movement

Let $\alpha = X^n$ and $\beta = Y^m$, where α dominates β. If β undergoes specifier movement and is merged with α, the resulting syntactic object is X^{n+1}, a projection of α.

Terminal nodes

Clause V: Terminal nodes are X^0s.

The following corollary is immediately derived from Clause V:

(25) Corollary I

If α is a phrase, then α is non-terminal and therefore has daughters.

Proof: Suppose that α is a phrase and at the same time is a terminal. Then, α would be a head (= X^0) by definition, but this contradicts the assumption that α is a phrase (= X^n, where n > 0). QED.

Let us now define domination and projection. To that end, let us first define n-degree daughter as follows:

(26) Definition

(i) B is a 1-degree daughter of A if and only if B is formed by External or Internal Merge of A and another syntactic object.

(ii) If B is an n-degree daughter of A and C is an m-degree daughter of B, then C is an (n+m)-degree daughter of A.

(iii) A itself is the only 0-degree daughter of A.

With this, domination is defined as follows:

(27) Definition

α **dominates** β if and only if there is a natural number n > 0 such that β is an n-degree daughter of α.

In order that maximal projections may include terminal nodes as well, which is a necessary assumption as I discuss in the next section, we have to guarantee that α itself is a projection of α. Therefore, I define projection in a recursive manner as follows:

(28) Definition

P is a **projection** of α, if

(i) P = α, or

(ii) P has a daughter D such that P is a projection of D as defined by Phrase Structure above and that D itself is a projection of α.

Nothing else is a projection of α.

Now I prove the following lemma:

(29) Lemma I

For every node α, there exists exactly one sequence $\{p_n\}$, such that $\alpha = p_0$ and for each $k > 0$, p_{k+1}, if it exists, is p_k's mother, and that all and only the members of $\{p_n\}$ are projections of α. We call $\{p_n\}$ α's projection sequence. When $\{p_n\}$ has more than one member, for any p_i and p_j, where $i > j$, p_i is a projection of p_j and p_i dominates p_j.

Proof: We construct $\{p_n\}$ with the following procedure:

1. Let $p_0 = \alpha$ and $k = 0$.
2. If p_k has a mother β such that β is a projection of p_k as defined by Phrase Structure, then let $p_{k+1} = \beta$. Otherwise, p_{k+1} does not exist.
3. If p_{k+1} does not exist, end this procedure.
4. Increase k by 1 and go to 2.

With this, we can construct a unique sequence $\{p_n\}$ for a given α. From the algorithm, it is clear that every member of $\{p_n\}$ is a projection of $p_0 = \alpha$. At the same time, the algorithm exhausts all the projections of α.

Now, suppose that p_i and p_j belong to $\{p_n\}$, where $i > j$. Clearly, p_j is an (i-j)-degree daughter of p_i and $i - j > 0$, so p_i dominates p_j. Let us now see that p_i is a projection of p_j. p_j determines a unique projection sequence starting with it. Let us call it $\{q_m\}$. Then, obviously $q_0 = p_j$, $q_1 = p_{j+1}$, …, and so $q_k = p_{j+k}$ for $k \geq 0$. Thus, p_i belongs to $\{q_m\}$ as well, and since every member of $\{q_m\}$ is a projection of $q_0 = p_j$, p_i is a projection of p_j. As the choice of p_i and p_j was arbitrary, it is shown that for any p_i and p_j, where $i > j$, p_i is a projection of p_j and p_i dominates p_j. QED.

Now we define maximal projection as follows:

(30) Definition

P is a **maximal projection** if and only if P does not have a mother that is a projection of P.

Thus, even terminal nodes may be maximal projections so long as they do not have a mother that is a projection of them. The following lemma is immediately derived:

(31) Lemma II

If α is a maximal projection, there is no projection of α except for α itself.

Proof: Suppose to the contrary that α is a maximal projection and that there exists a β such that β is a projection of α and $\beta \neq \alpha$. By Lemma I, α has a unique projection sequence $\{p_n\}$, where $p_0 = \alpha$. Since β is a projection of α, β belongs to $\{p_n\}$, and since $\beta \neq \alpha$, $\beta = p_i$ for some $i \geq 1$. Since p_i exists, it is evident from the construction of $\{p_n\}$ that p_k exists for all $0 < k < i$. Particularly, p_1 exists. In other words, α has a mother that is a projection of α. This contradicts the assumption that α is a maximal projection. QED.

Now, a brief note on terminology is in order. In the conventional model, A and A-bar movement are called phrasal movement. What is referred to as phrasal here are actually maximal projections, however. Chomsky (1994, 1995) proposes that clitics are heads and maximal projections at the same time, and since they are maximal projections they may undergo A movement. Although clitics may be maximal projections, according to Phrase Structure above, they should not be phrases unless there is some internal syntactic structure in them. Then, the relevant characteristic that distinguishes head movement from "phrasal" movement in the conventional model should be that while the "phrasal" movement moves maximal projections, head movement moves heads, which are not maximal projections. In my model, by contrast, what head movement moves is a head, but is at the same time a maximal projection. Thus, the distinction in the conventional model in this respect is lost in my model. Both kinds of movement move maximal projections. I discuss this point in the next section and in fact propose as an axiom.

What, then, distinguishes so-called "phrasal" movement from head movement in my model? One difference is that head movement moves heads exclusively while "phrasal" movement may move phrases. However, note that "phrasal" movement also moves heads such as clitics. Thus, when we see a movement of a phrase, we can tell that it is a "phrasal" movement, but a movement of a head is not necessarily a head movement. A clearer distinction between head movement and "phrasal" movement consists in that in the former, the moved element becomes a head of the newly created node, whereas in the latter, the moved element becomes a specifier of the newly created node. Thus, I use the term **specifier movement** to refer to so-called "phrasal" movement.

Now that our Phrase Structure is given, we want to make sure that the desired properties of syntax as discussed so far are derived as theorems. First, let us confirm that the following obvious lemma holds.

(32) Lemma III

If α is a head, then every node that α dominates is a head.

Proof: If α is a terminal node, there is no node that α dominates, and so the proposition vacuously holds. Thus, let us assume that α is not a terminal node.

Let P(n) be the proposition: every n-degree daughter of α is a head. We show that P(n) holds for any $n \geq 0$ by induction on n.

(i) P(0) holds, since α is α's only 0-degree daughter, and α is a head by assumption.

(ii) Suppose P(k) holds. If α does not have any (k+1)-degree daughters, P(k+1) vacuously holds. Therefore, assume that α has (k+1)-degree daughters, and let β be one of them. Since each node is binary branching, β has exactly one sister. Let us call it γ and let δ be the mother of β and γ. Then, since δ is β's mother, δ is a k-degree daughter of α, and since P(k) holds, δ is a head. γ is a (k+1)-degree daughter of α, since γ is δ's daughter. Then, β and γ must not have been merged by Internal Merge. If so, by Clauses III and IV of Phrase Structure, δ would have been a phrase, but δ is a head. This is a contradiction. Therefore, β and γ must have been merged by External Merge. If either of β and γ were a phrase, by Clause II, δ would have been a phrase, again a contradiction. Thus, both β and

γ are heads. Since β is an arbitrary (k+1)-degree daughter of α, it has been shown that if α has (k+1)-degree daughters, they are all heads. Hence P(k+1) holds.

By (i) and (ii), P(n) holds for any $n \geq 0$. Since every node that α dominates is an n-degree daughter of α for some $n > 0$, we have proven that every node that α dominates is a head. QED.

Now, we go on to prove the following related theorem, again an expected result:

(33) Theorem I
α is a phrase if and only if either α is formed by Internal Merge or α dominates a node that is formed by Internal Merge.

Proof: We prove the $\Rightarrow$ proposition (♠) and the $\Leftarrow$ proposition (♣) below.
(♠) α is a phrase $\Rightarrow$ either α is formed by Internal Merge or α dominates a node that is formed by Internal Merge.

For the purpose of obtaining a contradiction, assume that α is a phrase, and neither α is formed by Internal Merge nor α dominates a node that is formed by Internal Merge. In other words, α is a phrase which has been constructed only via External Merge.

Now, let Q(n) be the proposition: if α has an n-degree daughter, then α has at least one n-degree daughter that is a phrase. We prove that Q(n) holds for every $n \geq 0$ by induction on n.

(i) Q(0) holds since α is α's only 0-degree daughter and α is a phrase by assumption.

(ii) Suppose that Q(k) holds. If α has no (k+1)-degree daughter, then Q(k+1) vacuously holds. Now, assume that α has a (k+1)-degree daughter. Then, since its mother is a k-degree daughter of α, by Q(k), α has at least one k-degree daughter that is a phrase. Let us call it β. By Corollary I, β has daughters, and since β is a k-degree daughter of α, β's daughters are (k+1)-degree daughters of α. Then, at least one of its daughters must be a phrase, since if both its daughters were heads, by Clause I, β would be a head, and this contradicts the assumption that β is a phrase. Thus, α has at least one (k+1)-degree daughter that is a phrase. Hence Q(k+1) holds.

By (i) and (ii), it is shown that Q(n) holds for every $n \geq 0$. Now, let $N = \text{Max}\{n \mid \alpha \text{ has an n-degree daughter}\}$. Since Q(N) holds, α has an N-degree daughter that is a phrase. Let us call it γ. By Corollary I, γ has daughters, and since γ is an N-degree daughter of α, this means that α has (N+1)-degree daughters. However, this contradicts the definition of N.

(♣) Either α is formed by Internal Merge or α dominates a node that is formed by Internal Merge $\Rightarrow$ α is a phrase.

If α is formed by Internal Merge, by Clauses III and IV of Phrase Structure, α is a phrase.

Now assume that α dominates a node that is formed by Internal Merge. By Clauses III and IV, this node is a phrase. Then, it follows from the contraposition of Lemma III that α is a phrase. This completes the proof of the theorem. QED.

From Theorem II, the following corollary is obtained:

(34) Corollary II
α is a head if and only if it is terminal or is formed only by means of External Merge.

4.4 Head movement

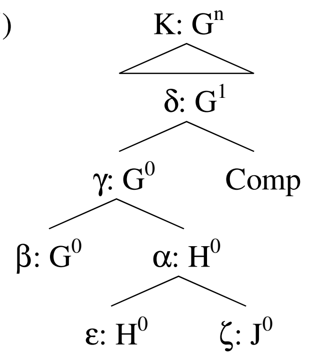
4.4.1 Locality revisited

Let us return to the locality in head movement that we observed in Section 4.1, repeated below:

(4) Locality in head movement
Given K, if there are nodes α, β and γ such that α and β are the daughters of γ, and β and γ are both heads of K, then only α is local to K with respect to head movement.

Consider the tree below:

(35)



According to (4), only α is local to K and so is able to undergo head movement in (35). This correctly captures the generalization, but the relation of α to K is not clear in an obvious way. In (35), β, γ, ε and ζ are heads, too, but they are not local to K according to (4), and so may not undergo head movement. What is the difference between α and all the other heads?

In the conventional model, what corresponds to γ is a terminal node (before head movement) and so does not have an internal syntactic structure, and it is what is local to K. This locality condition couched in the conventional model makes sense, since what is local to K is the head of K, which is an obvious relation.

Returning to (35), on the other hand, what is local to K is only α, but α is not even a head of K, nor is it the complement of the a head of K. Thus, α does not seem to bear any intuitively obvious relation to K. This suggests that (4) is rather a stipulation that captures the generalization of head movement, and that the locality itself may have to be defined in simpler terms.

To sum up, what is local to K in the conventional model corresponds to γ in (35), but we need (at least) α to be local to K in my model. Then, a simple and straightforward proposal is to say that γ and all nodes that it dominates are local to K. To formalize this, let us first define head of a given node K and the maximal head of K as follows:

(36) a. Definition of a head of K
For a given node K, H is a **head of** K, if and only if H is a head (= X^0) and K is a projection of H.

b. Definition of the maximal head of K
For a given node K, M is the **maximal head** of K if and only if M is a head of K and M does not have a mother that is a head. When M is the maximal head of K, we write M = H(K).

In (35), both β and γ are heads of K. β is not the maximal head of K, since β has a mother that is a head (viz. γ). On the other hand, γ is the maximal head of K, since it has a mother (viz. δ), but the mother is not a head (it is G^1). Since we want to say that γ and all it dominates are local to K, the locality can be defined recursively as follows:

(37) Locality in head movement
(i) For a given node K, H(K) is local to K with respect to head movement.

(ii) If X is local to K with respect to head movement, so are its sisters.

(37) is more desirable than (4) in the sense that it is defined in terms of an obvious relation to K, namely, the maximal head of K.

The locality in (20) is less stringent than (4), however. Although both predict α to be local to K in (35), (37) says that β, γ, ε and ζ are all local to K, while (4) says these are not. Thus, in addition to α, (37) predicts that all the heads of K and, if α is non-terminal, all nodes that α dominates are local to K. (4) is stringent enough to prevent these nodes from undergoing head movement, but now that we have replaced it with the less stringent (20), head movement of these nodes must be excluded not due to locality, but for some other independent reasons. We discuss this in the following subsections.

Now, in the remaining of this subsection, I would like to confirm that (part of) Head Movement Constraint is actually derived from the locality given in (37).

(38) Theorem II
If α undergoes head movement and is merged with K, there is no phrasal maximal projection L such that K dominates L and L dominates α.

Proof: For the purpose of obtaining a contradiction, suppose that there is a phrasal maximal projection L such that K dominates L and L dominates α.

L does not dominate H(K). To see this, assume to the contrary that L dominates H(K). By Lemma I, H(K) has a unique projection sequence $\{p_n\}$ such that p_0 = H(K). Since K is a projection of H(K), K = p_i for some i. Since any given node may have only one node that dominates it, {p_n} represents the one and only path in the tree from H(K) up through K. Then, since K dominates L and L dominates H(K), it follows that L belongs to {p_n}, i.e., L = p_j for some $0 < j < i$. Thus, K is a projection of L that dominates L. However, since L is a maximal projection, it follows from Lemma II that there is no projection of L except for L itself. This is a contradiction.

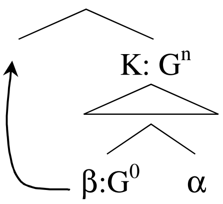
Now, since α undergoes head movement and is merged with K, α is local to K and so either α = H(K) or α is dominated by H(K). α ≠ H(K), because L dominates α by assumption but we have shown above that L does not dominate H(K). Therefore, α is dominated by H(K). Since L dominates α by assumption, both H(K) and L dominate α. Then, since L ≠ H(K) as L is a phrase by assumption and H(K) is a head, either H(K) or L dominates the other, but as we have seen above, L does not dominate H(K). Therefore, it must be the case that H(K) dominates L. Then, it follows from Lemma III that L is a head. This contradicts the assumption that L is a phrase. QED.

Theorem II and the adjacency condition discussed in Section 2.3 together constitute Head Movement Constraint.

4.4.2 Axiom of Movement

Here, we consider how head movement of heads of K is prohibited. This leads us to return to self adjunction. In (21), β is either the maximal head of K or dominated by the maximal head of K:

(39)



Therefore, β is local to K according to (37) and so should be able to undergo head movement, but that would yield a case of self adjunction. This can be excluded by simply assuming that only maximal projections may move at all. In (39), although β is a head (G^0), it is not a maximal projection and therefore is not eligible for head movement. The maximal projection of G is K = G^n above, but it may not move by any means since it is the very root of the tree. Moreover, G^n is not a head, so it is not eligible for head movement anyway. On the other hand, α may undergo head movement if it is a head, since it is a maximal projection.

In the conventional model, head movement is different from phrasal movement in that it is not movement of maximal projections. Not that we have assumed that head movement is also movement of maximal projections, we can generalize that syntactic movement is always movement of a maximal projection. I take this as an axiom in syntax:

(40) Axiom of Movement
Only maximal projections may move.

With this, we can derive the following corollary, which tells us that self adjunction is impossible:

(41) Corollary III
If α undergoes head movement and is merged with K, then, α is not a head of K.

Proof: Suppose that α is a head of K. Since α is moved and is merged with K, K $\neq \alpha$. Thus, K is a projection of α that is not α itself. By Axiom of Movement, however, α is a maximal projection, and therefore, by Lemma II, there exists no projection of α except for α itself. This is a contradiction. QED.

Let us now see that we can derive the fact that complements and specifiers do not undergo head movement.

(42) Theorem III
If α undergoes head movement and is merged with K, then α has a sister β and a mother γ such that β and γ are both heads and γ is a projection of β.

Proof: First, let us see that if α undergoes head movement, it always has a mother and a sister. Otherwise, α would be the root node, but by definition, the root node cannot undergo movement. Let us call α's sister β and α' mother γ.

Now, we show that γ is a projection of β. Since α moves, by Axiom of Movement, it follows that α is a maximal projection. Since α is a maximal projection, γ is not a projection of α. Hence γ must be a projection of β.

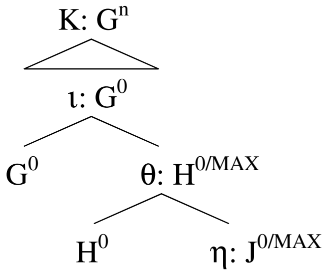
Finally, we show that β and γ are both heads, and this completes the proof. Since α undergoes head movement, α is local to K, and so either α = H(K) or α is dominated by H(K). Now, if α = H(K), then K would be a projection of α and therefore α would not be a maximal projection. However, since α moves, by Axiom of Movement, α is a maximal projection. This is a contradiction. Thus, we conclude that $\alpha \neq$ H(K). Then, α must be dominated by H(K), and therefore, either γ = H(K) or γ is dominated by H(K). If γ = H(K), then γ is a head, and since γ dominates β, by Lemma III, β must also be a head. If γ is dominated by H(K), so is β, since γ dominates β. Therefore, by Lemma III, β and γ must be both heads. QED.

According to Theorem III, if α undergoes head movement, both α's sister and mother are heads. Therefore, α cannot be a complement or a specifier.

4.4.3 Semantic types and binding

In Subsection 4.4.1, I noted two cases that need to be excluded for other reasons than locality. In the previous subsection, I discussed one of them, viz., self-adjunction. In this subsection, I discuss the remaining other case. Consider the tree below:

(43)



According to the locality given in (37), not only θ but also whatever θ dominates is local to K. By Lemma III, all the nodes that θ dominates are heads. Then, as long as they are maximal projections, they should be able to undergo head movement and be merged with K. Thus, η should be able to undergo head movement, and if η is non-terminal, any node that η dominates that is a maximal projection should also be eligible for head movement. However, in the derivation model that I have presented, when a head undergoes head movement and is merged with K, it is always the case that the moving head is the sister of a head of K. Thus,

(44) If α undergoes head movement and is merged with K, α is the sister of a head of K.

If this is correct, only θ and not η should be able to undergo head movement in (43). How can we derive this with our new locality in (27)?

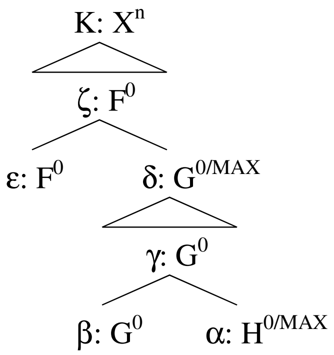
Now, let us suppose to the contrary that (44) is not correct. Let $\alpha = H^0$, and suppose that α undergoes head movement and is merged with K but α is not the sister of a head of K. By Theorem III, α has a sister β and a mother γ and if $\beta = G^0$, $\gamma = G^0$. Let δ be the maximal projection of β (δ might be identical to γ). By Lemma I, β has a unique projection sequence $\{p_n\}$, where $p_0 = \beta$. Since δ is a projection of β, δ belongs to $\{p_n\}$. Thus, $\delta = p_i$ for some i. p_i is the last member of $\{p_n\}$, since otherwise, δ would have a projection other than δ, and by Lemma II, δ would not be a maximal projection, but this contradicts the assumption that δ is a maximal projection. Since p_i is the last member of $\{p_n\}$ and at least p_1 (i.e., γ) exists, $i \geq 1$. Thus, δ (= p_i) dominates β (= p_0). Since α is β's sister, δ dominates α, too.

Now, it follows from the locality of α and Corollary III that α is dominated by H(K). Thus, both H(K) and δ dominate α, and therefore, it must be the case that either H(K) = δ or either of H(K) and δ dominates the other. However, if either H(K) = δ or H(K) were dominated by δ, then H(K) should belong to $\{p_n\}$ and so would be a projection of β. Then, K, which is a projection of H(K), would have to be a projection of β, but this contradicts the assumption that β is not a head of K. Therefore, we conclude that H(K) dominates δ.

Let δ's sister be ε and δ's mother ζ. Since H(K) dominates δ, H(K) also dominates ε, and by Lemma III, ε is a head. ζ may or may not be H(K). Either way, ζ is a head. This is obvious when ζ = H(K), and when ζ is dominated by H(K), it follows from Lemma III. Let $\varepsilon = F^0$. Since δ is a maximal projection, ζ is a projection of ε and therefore $\varepsilon = F^0$.

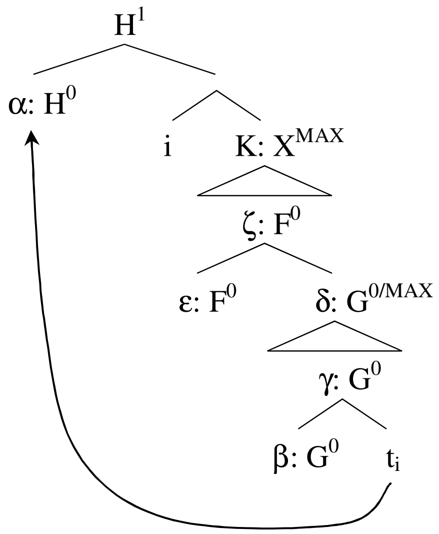
Thus, letting K = X^n, the configuration before the movement would look like the following:

(45)



Here, it might be that X = F and also it might be that δ = γ. By assumption, α undergoes head movement and is merged with K:

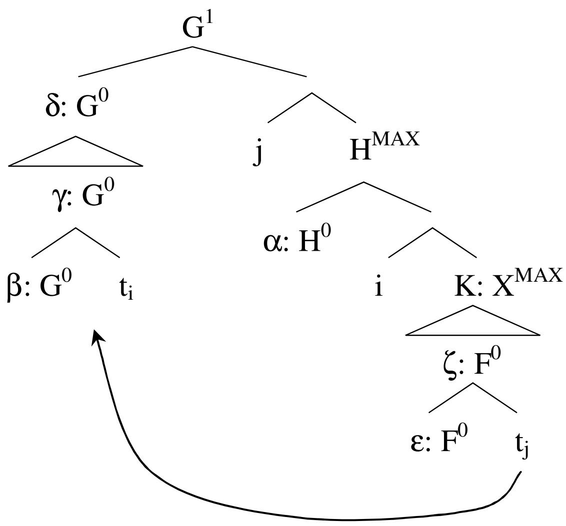
(46)



Here, a question arises as to whether the structure in (46) is interpretable. Assuming that (14) is correct, ε's denotation is of type <σ,τ> for some simple type σ and some T-type τ.

Now, let us assume that lexical entries are either of a simple type or of T-type. Since simple types are not functions and T-types are functions that yield denotations of a T-type, as long as only Functional Application is employed, every non-terminal node would have a denotation of a T-type. Therefore, since δ in (46) is not a terminal node, it has a denotation of a T-type. Thus, δ is not of a simple type. Particularly, δ is not of type σ. Then, since ε's denotation is of type <σ,τ>, δ must be a quantifier over elements of type σ so that it will leave a trace of type σ after movement. However, moving δ by head movement would lead the trace of α (t_i) to be left unbound, since δ would be moved across above the already moved α as illustrated below:

(47)



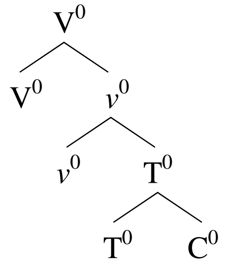
Thus, given that both α and δ have to move, in order to obtain an interpretable structure, δ should undergo head movement before α does, so that α moves out of the moved δ. (For the same reason, if X ≠ F, ζ has to move before δ does.) However, this contradicts the assumption that α is the one that undergoes head movement and is merged with K in (45). To sum up, as long as we assume that all lexical entries are either of a simple type or of a T-type, then (44) follows.

Unfortunately, although most lexical entries seem to be either of a simple type or of T-type, this assumption is wrong. Heim & Kratzer's (1998) analyzes the definite article (viz., *the*) as having a denotation of type <<e,t>,e>, but this is nether a simple type nor a T-type. Given that there are lexical items whose denotations are neither of a simple type nor of a T-type, (44) would be undermined. Suppose for example that in (46), α is a quantifier over elements of some simple type υ (possibly υ = σ) and so has left behind a trace of type υ, and β is a terminal node whose denotation is of type <υ,σ>, and δ = γ. Then, δ (= γ) would have a denotation of type σ, and thus the structure is interpretable and so δ would not have to move. Luckily, however, we do no know of any lexical items with denotations of type <υ,σ> for some simple types υ and σ. Therefore, as long as the definite article is the only exceptional item whose denotation is neither of a simple type nor of a T-type, (44) can be maintained.

4.6 Percus's problem

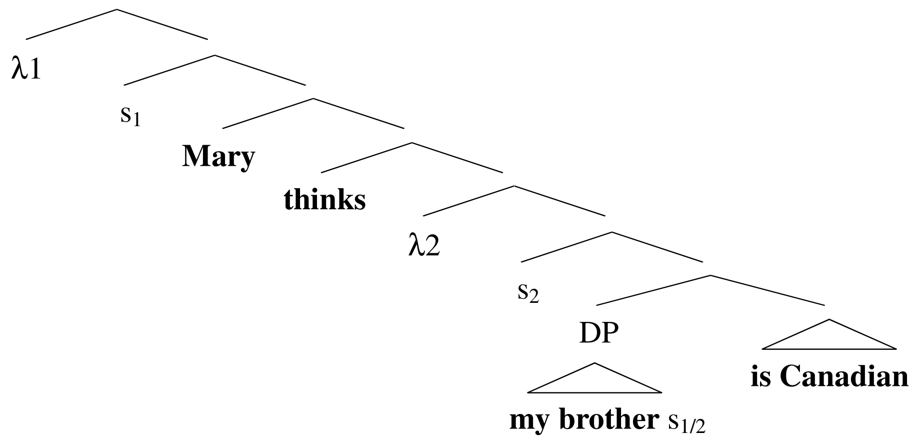
In this section, we finally come back to Percus's problem. For the sake of simplicity, let us ignore the S head that was introduced in Chapter 3. Recall that the S head was postulated for semantic reasons, and there is no particular syntactic evidence for the existence of such a head and that there is an alternative way in which T will act as a quantifier over spatiotemporal regions. Whether we have S in the tree or not does not affect the validity of the main points of this section's discussion. Thus, we assume that the verb of the main clause begins as follows:

(48)



Let us first return to the licit binding configuration we looked at in Section 1.2:

(49)



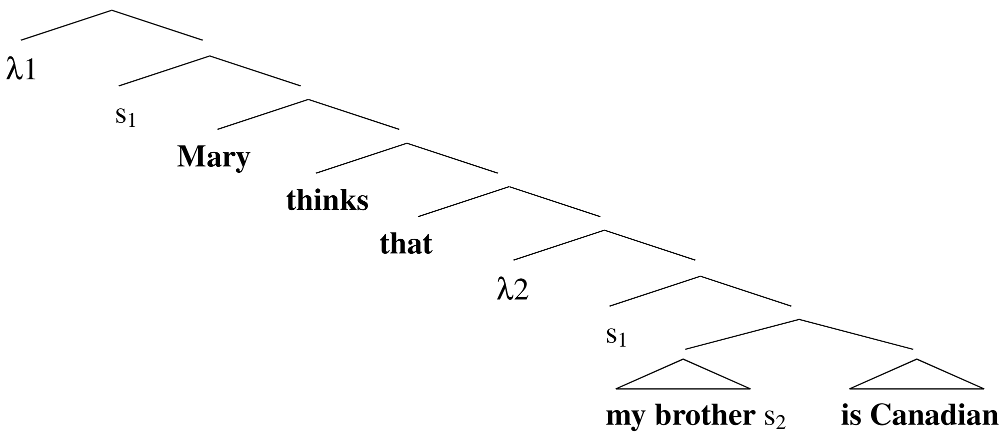
In my derivation model, the λ operators that bind world variables are introduced by head movement of C^0. That is, λ1 is introduced by the head movement of the (phonetically empty) maximal C^0 in the matrix clause, and λ2 by the head movement of the (phonetically empty) maximal C^0 in the embedded clause in the course of the derivation. Note that unlike the structures that Percus considers, in my model, the world variable in the VP of each clause should be outside the VP and the sisters of a T head in the same clause. For the similar problem regarding Tense, the λ operator that binds the time variable of the main predicate of a clause is introduced by head movement of the T^0 of the same clause.

Thus, as long as the derivation of a clause begins with the formation of the complex head such as in (48), we automatically obtain Percus's generalization. Since C^0 is part of the complex head for the main predicate of a clause, the world argument for the main predicate of the clause will automatically be bound by the λ right below the clause's complementizer, which is the C^0 that has undergone head movement in the clause. The same argument holds for the Tense. Since T^0 is part of the complex head for the main predicate of a clause, the time argument for the main predicate of the clause will automatically be bound by the λ right below the clause's Tense, which is the T^0 that has undergone head movement in the same clause.

If the λ created by the head movement of the embedded C^0 (indexed as 2 in (49)) binds the world argument of *my brother* in addition to the trace of the moved head in (49), a *de re* interpretation of *my brother* is obtained. On the other hand, if the world argument of *my brother* is bound by the λ created by the head movement of the matrix C^0 (indexed as 1), a *de dicto* interpretation of *my brother* is obtained. This could be achieved either by virtue of long distance binding or by moving the DP out of the embedded clause (see Section 3.4.3 for a relevant discussion).

By contrast, in an illicit binding configuration that violates Percus's generalization, the world argument for the predicate of the embedded clause is not bound by the nearest binder λ2 as in the following. Below is the configuration that we looked at in Section 1.2 except that I have added the overt complementizer *that* for expository purposes:

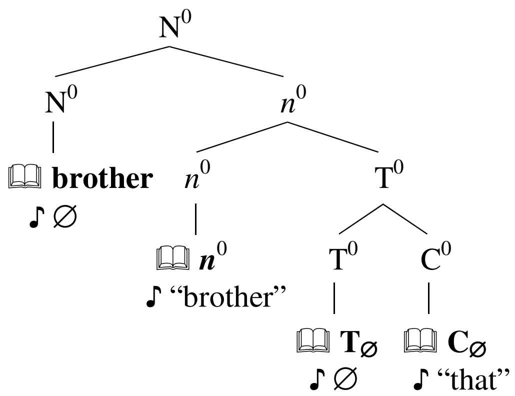
(50)



Here, the world argument for the predicate of the embedded clause is bound by λ1 in the matrix clause, and λ2 binds only the world argument inside the noun phrase *my brother*. In order to obtain such a binding configuration in my derivation model, the maximal C^0 in the embedded clause (i.e., the complementizer of the embedded clause) must move out of *my brother*. However, this is impossible since *my brother* is a specifier, and head movement is not possible out of a specifier as we have seen above (Theorem III). Such a head movement is impossible even if *my brother* itself has moved to some other position. Since *my brother* is a specifier, it may only undergo specifier movement, and therefore it will always be a specifier. Therefore, head movement from inside it will never be possible.

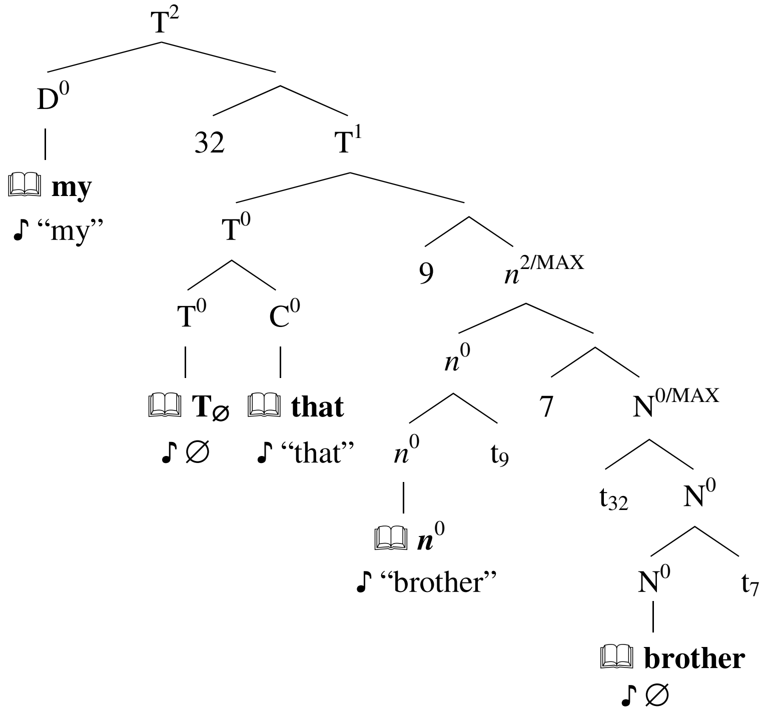
For more concrete exposition, let us see how (50) would be derived in my model in more detail. As I have just explained above, the complementizer of the embedded clause must originate inside the subject of the embedded clause. Thus, a head complex like (51) should be created for the subject of the embedded clause:

(51)



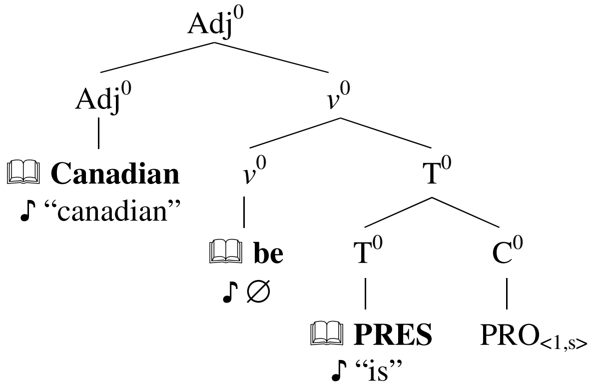
Let us assume that the possessive pronoun *my* is a D^0 and is merged as a specifier of N. After T^0 undergoes head movement, *my* moves to a specifier of T:

(52)



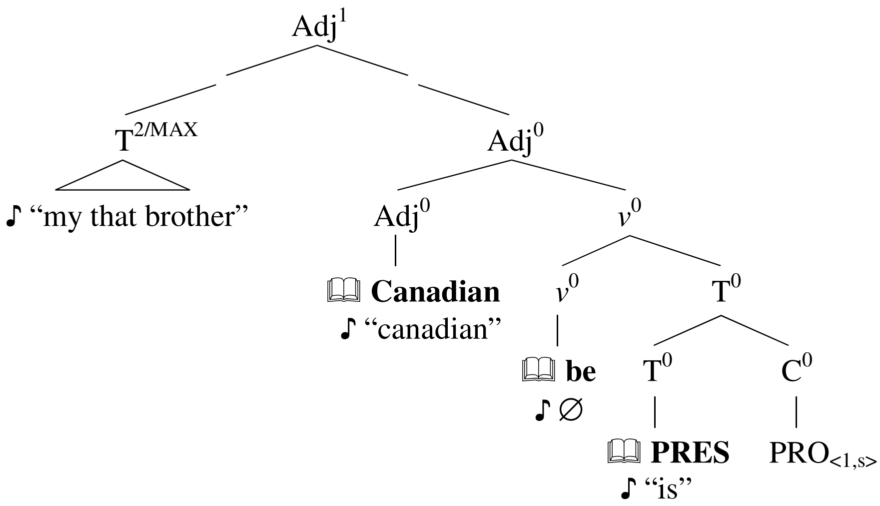
On the other hand, since the world variable of the main predicate of the embedded clause is to be bound by λ1 in the matrix clause, the complex head for the main predicate of the embedded clause involves a C^0 that is a PRO indexed as 1, which should eventually get bound when head movement of C^0 takes place in the matrix clause. Assuming that the copula *be* is a v^0, the complex head for the main predicate would look like the following:

(53)



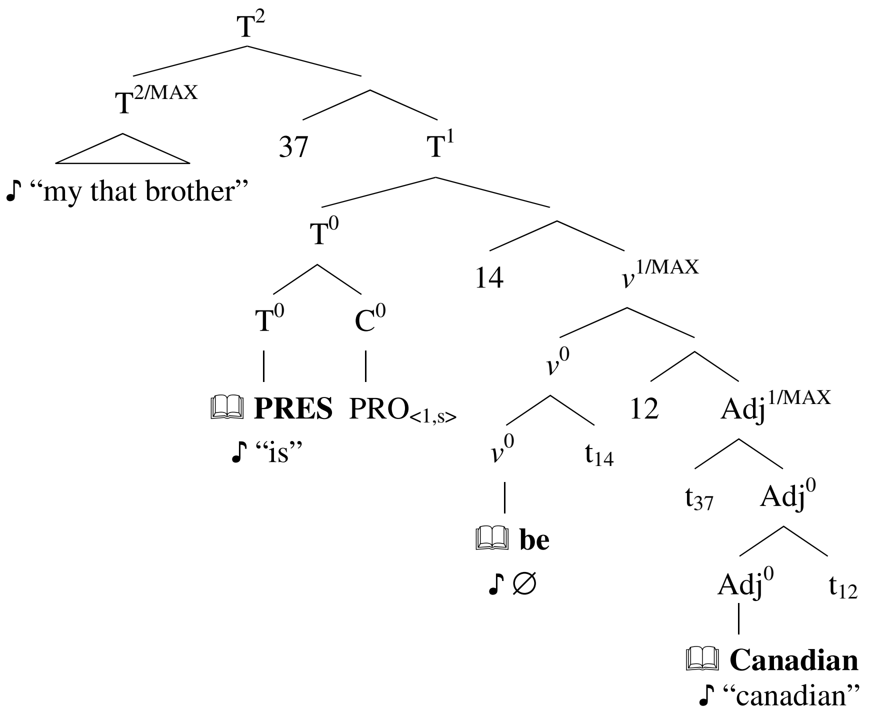
Now, we assume that the subject noun phrase created in (52) is merged with (53):

(54)



Here, the resultant node should be an Adj^2 as the generalization in (19) predicts. Thus, the subject noun phrase (T^2) becomes a specifier. Now, v^0 and T^0 undergo head movement and the subject noun phrase (T^2) would move to a specifier of T by A movement:

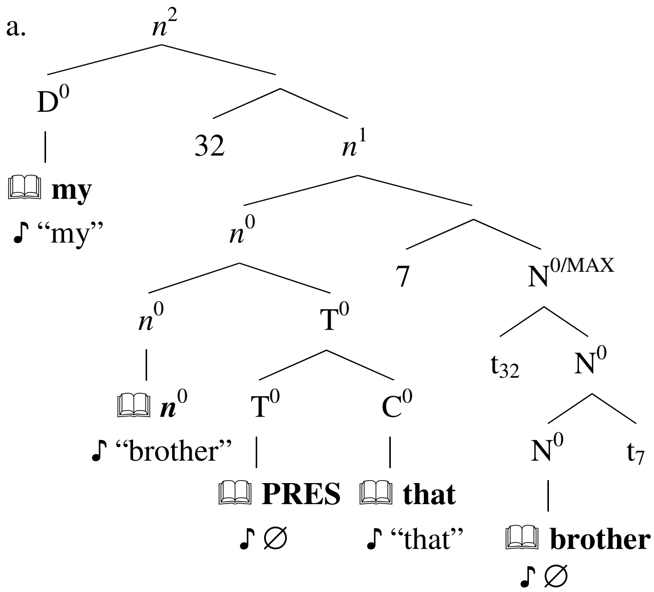
(55)



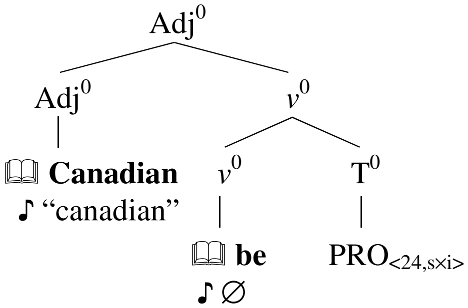
Again, the subject noun phrase (T^2) is a specifier. To achieve the binding configuration in (50), the C^0 inside the subject noun phrase should be moved by head movement and merged with the root node in (55). However, since the subject noun phrase is a specifier, head movement out of it is impossible.

Now, let us suppose that head movement of T does not take place inside the subject noun phrase, but T^0 undergoes head movement out of it. In other words, we are supposing that the Tense of the main predicate of the embedded clause originates inside the subject noun phrase, just like the complementizer of the embedded clause does. Let us further assume that the T^0 that takes part in the head complex for the main predicate of the embedded clauses is a PRO of type s×i which gets bound long-distance by the binder associated with the tense of the matrix clause. In this case, the derivation of the subject noun phrase stops at (56a) before it is merged with the main predicate shown in (56b):

(56) a.

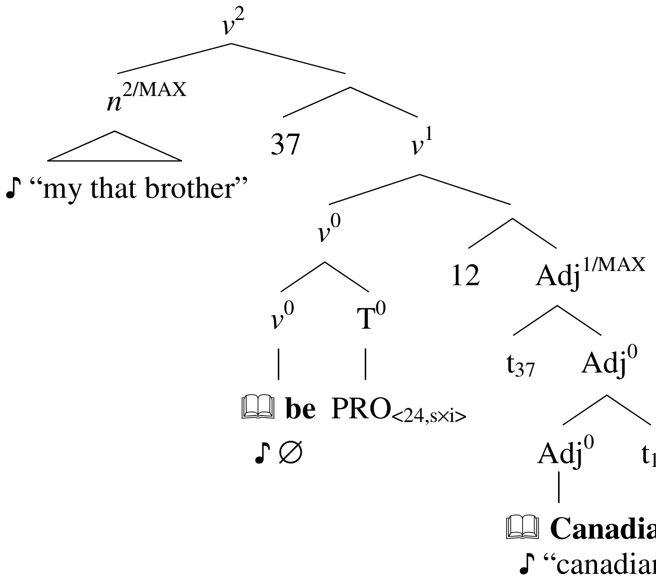


b.



The merger of (56a) and (56b) would yield an Adj^1 as the generalization (19) predicts. Then, after head movement of v^0, the subject noun phrase may undergo A movement to a specifier of v:

(57)

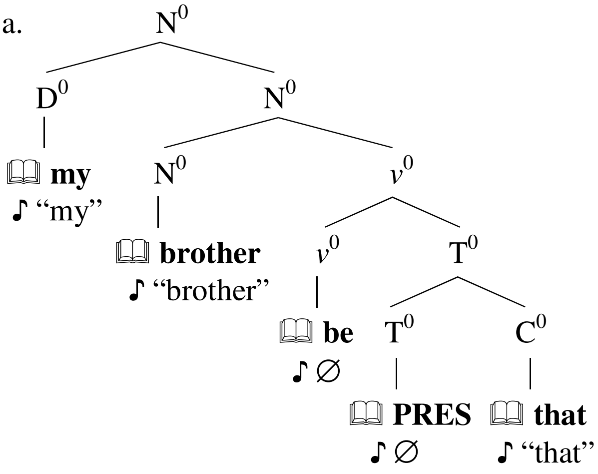


If the complex head T^0 inside the subject noun phrase undergoes head movement and is merged with the root node in (57), then the derivation can proceed to yield a binding configuration that exhibits both a world binding and a temporal binding that violate Percus’s generalization. Again, this is excluded, since the subject noun phrase is a specifier, head movement out of it is impossible.

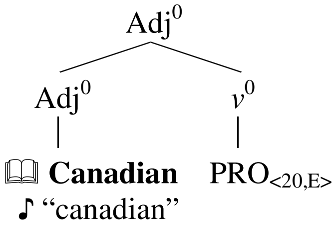
Let us recapitulate the discussion. According to the generalization in (19), the noun phrase (individual) arguments of a predicate are generated as specifiers of V, Adj, *v*, etc. Therefore, even if they move, they will always be specifiers, since specifiers may not undergo head movement, but only specifier movement (see Theorem III). Now, in order to obtain illicit binding configurations that violate Percus’s generalization, head movement should take place out of a noun phrase argument of the main predicate. However, this is never possible, since noun phrase arguments are always specifiers and head movement out of a specifier is impossible (Theorem III). This is how Percus’s generalization is derived in my framework.

However, there is actually one way around the problems I discussed above to achieve an illicit binding configuration that violates Percus’s generalization. Let us assume that the subject noun phrase begins with the head complex in (58), which has a v^0 *be* instead of an n^0, and that the main predicate begins with the head complex in (58b), where v^0 is an event PRO and there are no T^0 and C^0.

(58) a.

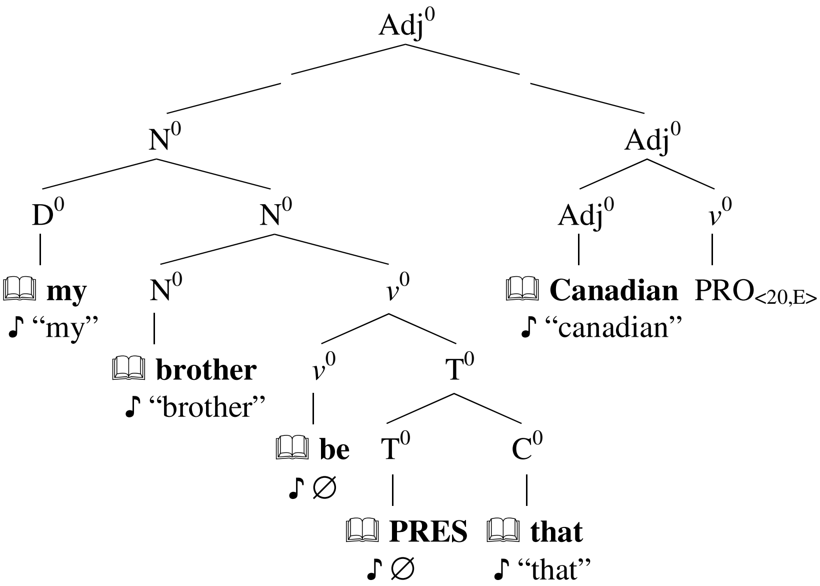


b.



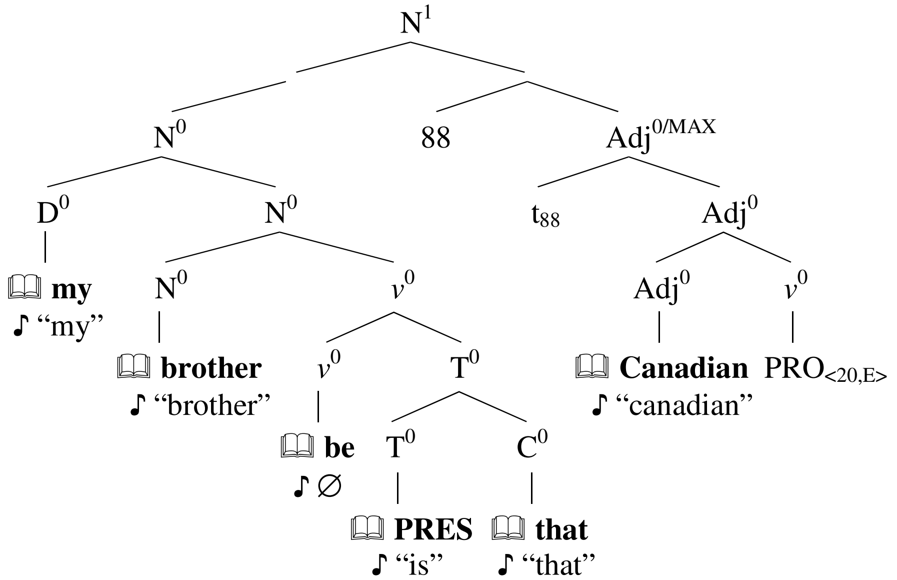
Suppose now that (58a) and (58b) are merged. According to the generalization in (19), the Adj^0 is projected just as in the cases we have considered above. Also, since this is an external merger of two heads, the resulting node is another head (Clause I of the Phrase Strucutre). Thus, the structure in (59) is obtained:

(59)



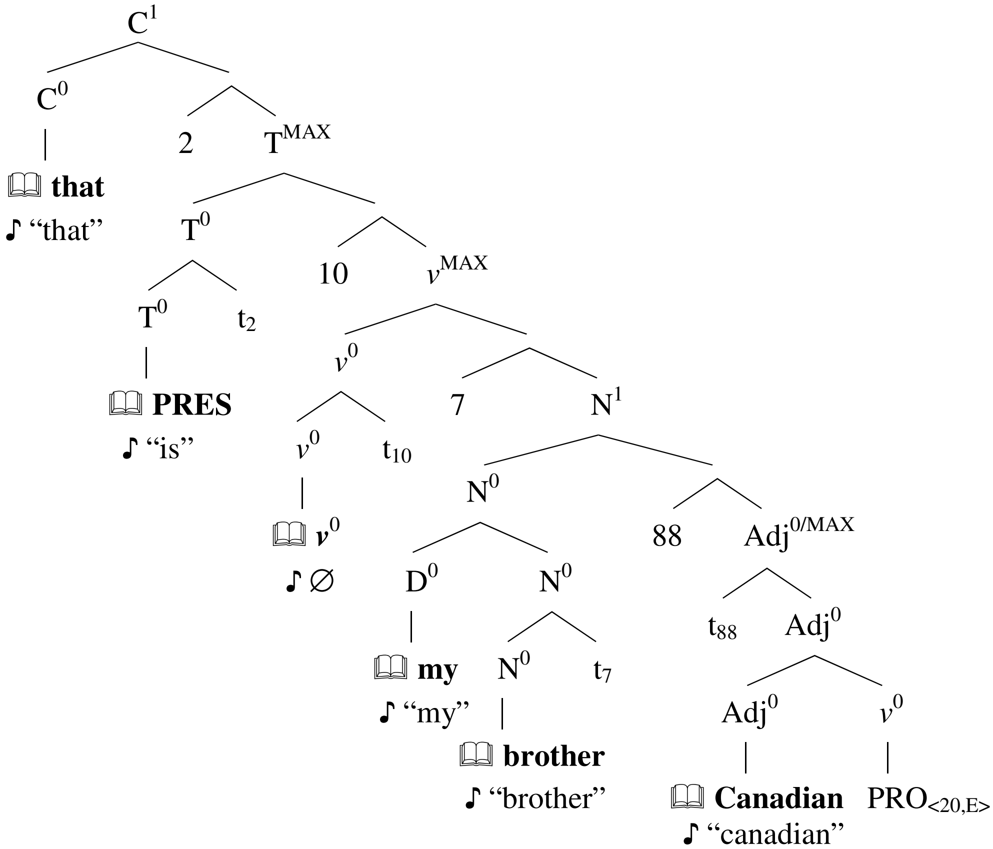
Here, the maximal head of the root node (Adj^0) is itself and it dominates the maximal N^0. Therefore, the maximal N^0 is local to the root node according to (36), and thus, the maximal N^0 is able to undergo head movement, which would result in the following:

(60)



Here, the maximal v^0 is local to the root node N^1 and so is eligible for head movement. Thus, the derivation can proceed to yield the embedded clause:

(61)



Now, if the event $PRO_{<20,E>}$ gets bound long-distance by the binder created by the head movement of v^0 in the matrix clause, a structure will be obtained where Percus's generalization is violated.

In order that such a case may not arise, I suggest that such head movement of N^0 as has occurred in (60) should be ruled out. In (59), the structure is already interpretable, even though the head movement of N^0 has not taken place. Furthermore, this head movement does not invert any scopes. Then, if we assume that head movement without semantic consequences should be ruled out for economical derivations, (60) will not obtain, and thus neither will (61). In addition, such derivations as this last case would not be tenable when the main predicate is inflected for Tense (e.g., tensed verbs). In such cases, a head in the predicate (e.g. V^0) and the T^0 will have to be morphologically merged. However, they do not start in the same head complex in this kind of derivation, and thus the morphological merger would be impossible, which would also mean that V to T movement seen in languages like French would be impossible. Assuming that morphology dictates whether or not some morphemes should be morphologically merged, failure to morphologically merge a predicate head such as V^0 and the T^0 to form an inflected form would lead either to a crashed derivation or to an unpronounceable sentence, depending on the assumption. Thus, in any event, we can safely conclude that derivations like (61) above should not arise.

Conclusion

In this paper, I have proposed a new model of head movement in order to solve two independent problems. The first problem is a syntactic property of the conventional model of head movement. In the conventional model, head movement merges a moving head with a non-root node of a tree, which operation results in a landing site that does not c-command its trace. This is fundamentally incompatible with the minimalist framework, in which only the root node is accessible. The second problem is how one can derive Percus's (2000) generalization that non-individual arguments of the main predicate of a clause must be bound by the nearest possible λ operators. These two problems have turned out to be connected, because in the proposed model, heads that undergo head movement are quantifiers over worlds, times, etc., that create bindings of non-individual arguments of the main predicate when they move.

In this derivation model, the derivation begins with forming a head complex and proceeds as heads move out of the head complex by head movement to remedy a semantic type mismatch. Therefore, the first problem does not arise since the merger of heads is done prior to the movement, and thus, the moved head always c-commands its trace. Also, Percus's generalization has turned out to be a direct consequence of the locality condition of head movement that says that head movement cannot move a head out of a specifier, since bindings of non-individual arguments of the main predicate of a clause are created by virtue of head movement.

In addition, I have explored how various word orders of different languages can be derived in my model in Chapter 2, and in Chapter 3, I have developed the appropriate semantics for my model and discussed the interpretations and structures of cardinal and presuppositional noun phrases. In Chapter 4, I have discussed that this new model of head movement naturally leads to a new phrase structure according to which Internal Merge plays an essential role in forming phrases (Theorem I).

Appendix

As I briefly discussed in Section 3.4, Musan (1995) presents a possible analysis of her Thesis of Temporal Dependence based on Diesing (1992), which can be referred to as the scope approach. According to Diesing, presuppositional noun phrases (including noun phrases with a weak determiner in their strong meaning) are outside the VP domain (which would correspond to *v*P in the recent literature) at LF, whereas cardinal noun phrases remain inside the VP domain. On the scope approach, there is some temporal quantifier just above VP that obligatorily binds the time arguments of nouns that are inside *v*P together with the time argument of the verb. Thus, cardinal noun phrases are necessarily temporally dependent, since they stay inside VP and their time arguments are obligatorily bound by the temporal quantifier. On the other hand, presuppositional noun phrases get scoped out of this temporal quantifier, and so their time arguments cannot be bound by the quantifier, and hence they obtain temporally independent readings.

Musan argues, however, that such an analysis is not tenable since it cannot account for certain sentences. She identifies the temporal quantifier above VP is a temporal adverb of quantification (TAQ), and assumes that the Tense is merely a restrictor to the quantification. On this approach, it seems difficult to account for the German sentence in (1a) under the reading described in (1b):

(1) a. In den sechziger Jahren spielten meistens alle Professoren, deren Eltern
in the sixties played mostly all professors whose parents

gerade in Urlaub waren, Federball.
just in vacation were badminton

"In the sixties, most often all professors whose parents were just then on vacation mostly played badminton."

b. For most times t, such that t is the past and in the sixties, all x such that x is now professors and x's parents were on vacation at t, x played badminton at t.

Here, the Tense and the temporal adverbial are restrictors for the adverb of quantification. Also, since the vacation time is bound by the adverb of quantification, the subject noun phrase containing the relative clause must be in the scope of the adverb of quantification. Thus, we would have something like the following at LF:

(2) meistens [$_{\text{RC(meistens)}}$ PAST & in-den-sechziger-Jahren]
[$_{\text{NS(meistens)}}$ alle [$_{\text{RC(alle)}}$ Professoren denen …] [$_{\text{NS(alle)}}$ Federball spielten]]

(RC: restrictive clause, NS: nuclear scope)

Given (2), the scope analysis predicts that *Professoren* should be temporally dependent, since it is inside the scope of the adverb of quantification and therefore its time argument must be obligatorily bound by it. Yet, *alle Professoren* is a presuppositional noun phrase and can have a temporally independent interpretation as in (1b). Thus, a scope paradox is obtained.

Since the scope analysis fails, Musan exploits stages in her semantics in order to explain the difference between the temporal interpretations of presuppositional and cardinal noun phrases. More precisely, she proposes that noun phrases contain quantification over stages of individuals, rather than quantification over individuals. For her, stages of an individual are (possibly disconnected) temporal slices of an individual. Presuppositional noun phrases have an extra quantification over stages of individuals that makes it possible for presuppositional noun phrases to be temporally independent of the main predicate. Cardinal noun phrases cannot be temporally independent, since they lack such an extra quantification over stages of individuals. (3a) is the denotation of the presuppositional determiner *most* (p. 115) by Musan (with very minor modifications), and (3b) is the denotation of the cardinal determiner *few* that I abstracted from her truth conditions.

(3) a. $[\![\textbf{most}]\!] = \text{f: } D_{et} \rightarrow D_{<et,<et,t>>}$

For any k, h, j $\in D_{et}$, $f(k)(h)(j) = 1$
iff for most maximal stages $x_{st} \in D_e$,
such that there are an $x'_{st} \in D_e$ and an $x''_{st} \in D_e$,
such that $x_{st} = x'_{st} \oplus x''_{st}$
and $k(x'_{st}) = 1$
and $h(x''_{st}) = 1$,
there is a stage $y_{st} \in D_e$,
such that y_{st} is part of $x_{st} \in D_e$
and $j(y_{st}) = 1$

b. $[\![\textbf{few}]\!] = \text{f: } D_{et} \rightarrow D_{<et,<et,t>>}$

For any k, h, j $\in D_{et}$, $f(k)(h)(j) = 1$
iff for few maximal stages $x_{st} \in D_e$,
such that $k(x_{st}) = 1$,
there is a stage $y_{st} \in D_e$,
such that y_{st} is part of $x_{st} \in D_e$
and $h(y_{st}) = 1$
and $j(y_{st}) = 1$

Here, the first argument k represents the resource domain, a silent variable which can be modified by the context, the Tense and temporal adverbials in the clause in which the noun phrase containing the determiner appears. The second argument h is the denotation of the noun that the determiner combines with, and the third argument j corresponds to the denotation of the main predicate. $\oplus$ is a mereological sum operation which fuses two stages of the same individual to yield a (possibly bigger) stage of the same individual. Also, stages in the second line are "maximal" so that different stages of one individual do not get counted separately.

As can be seen from the denotations, the difference between the presuppositional and cardinal determiners that the former (in a way) maps the noun phrase denotation to the restrictive clause, where as the latter maps it to the nuclear scope. Furthermore, in the restrictive clause, properties of individuals are combined with a non-Boolean "and" (i.e., a mereological sum operation on stages), whereas in the nuclear scope, Boolean "and" is utilized to intersect the denotations of properties. According to this analysis, the truth conditions of a sentence containing a cardinal noun phrase involve a part that claims existence of a stage of an individual that satisfies both the property denoted by the noun that the cardinal determiner attaches to and the property denoted by the main predicate. Since a stage is a temporal slice of an individual, this means the time of the noun and the time of the main predicate coincide. This is why cardinal noun phrases always have temporally dependent interpretations.

Although Musan's theory captures the semantics of presuppositional and cardinal noun phrases with an insightful way, it has some problems. First, in her theory, the temporal interpretation of the main predicate is given in an indirect way by virtue of taking a stage of an individual which has a certain temporal extension, and this process is mediated by the denotation of the determiner of a noun phrase that the main predicate combines with. That the predicate time is compatible with the Tense and adverbials in the clause is assured by the silent resource domain variable that the determiner takes being somehow indirectly modified by these. Obviously, this is rather a strange assumption of how main predicates are interpreted temporally. This problem becomes obvious when one considers sentences that do not have nouns with determiners:

(4) John sneezed.

In order to account for such an example, we would need to assume that proper names like *John* are also quantifiers over stages of individuals, which might seem like a stipulative complication. Even if this was not a problem, the following example would certainly be problematic.

(5) It rained.

Here, with *it* being an expletive, the predicate *rained* has no individual argument to take, but by assumption, without taking a stage of an individual, a predicate's time cannot be specified. Thus, it seems necessary to posit that the main predicate directly takes either a time argument or an event argument or something like that which has its own temporal extension. However, once we do this, Musan's explanation of why cardinal nouns have temporally dependent interpretations would lose its power, since then, without further stipulations, the determiner would no longer have control over the time of the main predicate.

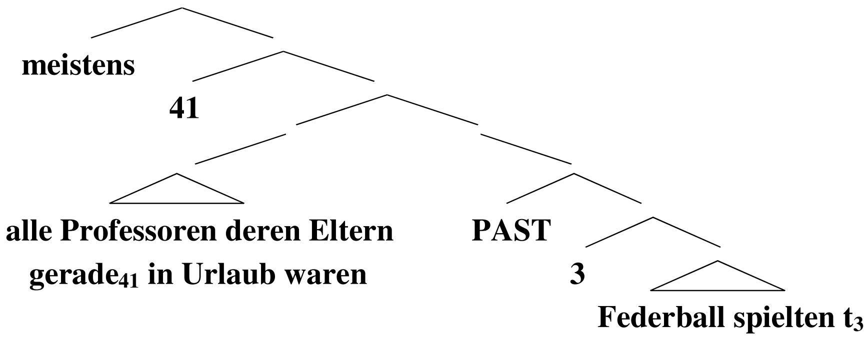
Secondly, Musan's theory has nothing to say about Diesing's findings that presuppositional noun phrases are outside VP (presumably in SpecTP) at LF and cardinal noun phrases inside VP. In Musan's theory, any noun phrases should be able to appear in either position at LF, since their time argument is never bound by anything outside them. Crucially, cardinal noun phrases should be able to appear in SpecTP, since in Musan's theory, it is not the case that the temporal interpretation of cardinal noun phrases is dependent on the temporal interpretation of the main predicate, but the temporal interpretation of the main predicate is dependent on the temporal interpretation on a cardinal noun phrase. On the other hand, the scope analysis based on Diesing's theory at least predicts that cardinal noun phrases can never appear outside the VP domain (assuming that whatever temporal quantifier is responsible for the temporal interpretation of the main predicate is right above VP). Musan says about the case we looked at in (1) and (2) above:

> Of course, one could avoid the contradiction by assuming a stipulation like "The TAQ binds time arguments of cardinal noun phrases in its scope obligatorily but not time arguments of presuppositional noun phrases in its scope". But if one adopts this stipulation, then one reaches a point where the distribution of temporally dependent and temporally independent noun phrase interpretations is attributed to some internal property of cardinal and presuppositional noun phrases and doesn't depend crucially on scope anymore (p. 92).

Apparently, "attributing the distribution of temporally dependent and temporally independent noun phrase interpretations to some internal property of cardinal and presuppositional noun phrases" is exactly what Musan is doing, and by giving up a possibility of a mixed explanation that would utilize both scope and noun internal properties, she has lost an account for the distribution of cardinal noun phrases that scope analysis has. Thus, although the scope approach by itself does not seem to offer an explanation of why presuppositional noun phrases should move out of VP, it explains why cardinal noun phrases remain inside VP, and this is an advantage over Musan's theory. In any case, that the scope approach does not predict the distribution of presuppositional noun phrases cannot be a reason to reject this approach completely, since neither does Musan's theory predict it.

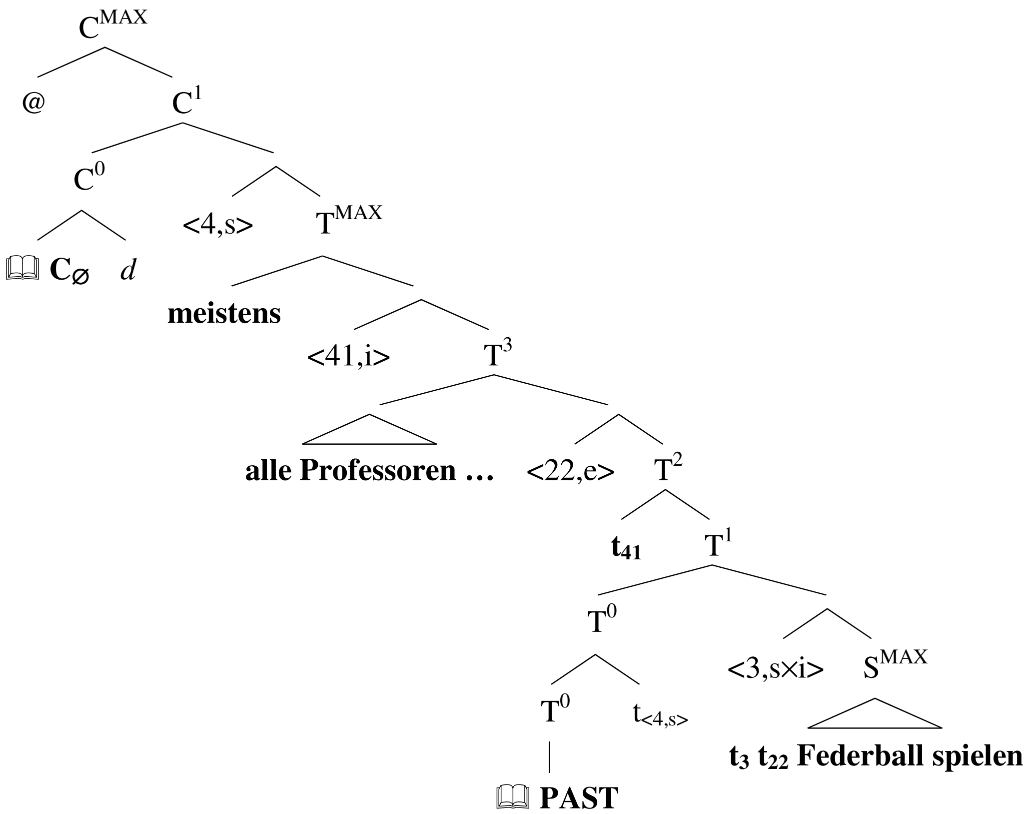
Third, it is actually not so obvious that a scope analysis has to run into a paradox in analyzing the example in (1) as Musan describes, since this is merely a consequence of her assumption that temporal adverbs of quantification are the only temporal quantifiers and that Tense is only a restrictor to them. Once we assume that the Tense is a temporal quantifier, a scope paradox need not arise. On such an approach, the presuppositional subject noun phrase in (1) *alle Professoren, deren Eltern gerade in Urlaub waren* need only be outside the scope of the Tense (i.e., at SpecTP), and not outside the scope of the temporal adverb of quantification (*meistens*). This way, *Professoren* receives a temporally independent interpretation. Since the Tense and the temporal adverb of quantification are different quantifiers, they each take their own scope, and thus their scopes do not have to coincide. Thus, a time argument in the relative clause *deren Eltern gerade in Urlaub waren* may in principle be bound by the temporal adverb of quantifier while the whole subject noun phrase is outside the scope of the Tense:

(6)



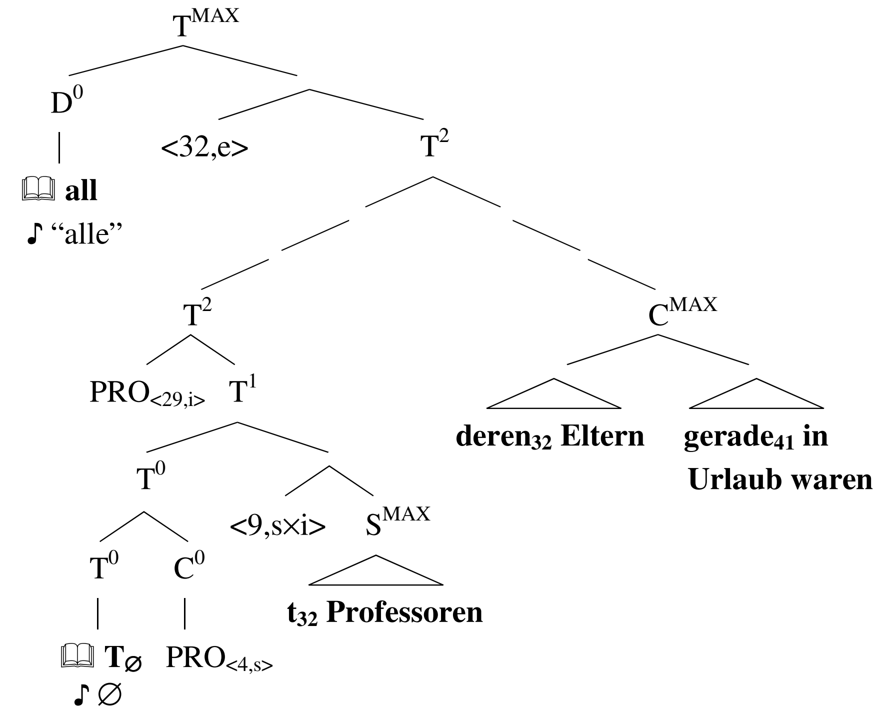
In my derivation model, this binding configuration can be achieved by the following tree:

(7)



Here, I am assuming that *meistens* has originated in a SpecT and moved to c-command the subject noun phrase, leaving a trace that acts as the resource domain of the temporal quantifier denoted by the T head. The structure of the subject noun phrase looks like the following:

(8)



Here, I am assuming that the relative clause is adjoined to T^2 of this presuppositional noun phrase, and the relative pronoun is bound by the binder created by the movement of the determiner *alle*. Although I cannot discuss the semantics of the relative clause in detail here, I simply assume that gerade gets bound by the binder associated with *meistens*, and provides the exact time period in which parents of professors were on vacation.

Let us now compute the semantics of (7). Under variable assignment a, the denotation of (8) is computed to be something like (9a), and the denotation of the sister of *meistens* is computed to be something like (9b):

(9) a. $\lambda g \in D_{et}.\ \forall x\ [\exists s \in D_{s\times i\times p}$: s is a situation in now [x is a professor in s] $\wedge$
on-vacation(x's parents)(a(41)) $\rightarrow$ g(x)]

b. $\lambda t \in D_i.\ \forall x\ [\exists s \in D_{s\times i\times p}$: s is a situation in now [x is a professor in s] $\wedge$
On-vacation(x's parents)(a(41))
$\rightarrow \exists t'$ [t' < now $\wedge$ t' $\subseteq$ t $\wedge$ play-badminton(x)(t')]]

To be more precise, play-badminton(x)(t') = 1 if and only if there is a situation s' whose time is t' such that x plays badminton in s'. Now, let us give the following denotation to *meistens*:

(10) $[\![\textbf{meistens}]\!]^{c,a} = \lambda f \in D_{<i,t>}.\ \mathrm{MOST}(t \in D_i)\ [f(t)]$

If we apply (10) to (9b), we obtain a truth condition that would be true in almost any scenarios, since there should be many times when no parents of current professor was on vacation, and such times vacuously satisfy (9b). Thus, we should modify (9b) with a presupposition as follows:

(11) $\lambda t \in D_i$: $\exists x$ [$\exists s \in D_{s \times i \times p}$: s is a situation in now [x is a professor in s] $\wedge$
On-vacation(x's parents)(a(41))]
$\forall x$ [$\exists s \in D_{s \times i \times p}$: s is a situation in now [x is a professor in s] $\wedge$
On-vacation(x's parents)(a(41))
$\rightarrow \exists t'$ [$t' <$ now $\wedge$ $t' \subseteq t$ $\wedge$ play-badminton(x)(t')]]

Then, the truth condition of (7) would be roughly like the following:

(12) MOST($t \in D_i$: $\exists x$ [$\exists s \in D_{s \times i \times p}$: s is a situation in now [x is a professor in s] $\wedge$
On-vacation(x's parents)(a(41))])
[$\forall x$ [$\exists s \in D_{s \times i \times p}$: s is a situation in now [x is a professor in s] $\wedge$
On-vacation(x's parents)(a(41))
$\rightarrow \exists t'$ [$t' <$ now $\wedge$ $t' \subseteq t$ $\wedge$ play-badminton(x)(t')]]]

This is an appropriate truth condition for (7). Note that in order for (7) to be true, professors do not have to be doing badminton throughout the period in which their parents were on vacation, but they need to play badminton only once during their parents' vacation. This is captured by $t' \subseteq t$ in (12).

References

Abusch, Dorit. (1994). The sequence of tense revisited: Two semantic accounts of tense in intenstional contexts. In Hans Kamp (ed.), *Ellipsis, Tense, and Questions. Dyana Delivarable* R2.2.B, ILLC/Dept. of Philosophy, University of Amsterdam, 1994, pp. 85-139.

Bayer, Josef. (1996). *Directionality and logical form: On the scope of focusing particles and wh-in-situ*. Kluwer Academic Publishers.

Chomsky, Noam. (1994). Bare phrase structure. MIT occasional papers in linguistics 5, MIT. Reprinted In Webelhuth (ed.), *Government and binding theory and the minimalist program*. Oxford: Blackwell 1995.

Chomsky, Noam. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. (2001). Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*. Cambridge, MA: MIT Press.

Diesing, Molly. (1992). *Indefinites*. Cambridge, MA: MIT Press.

Den Dikken, Marcel, Richard Larson & Peter Ludlow (1997). Intensional transitive verbs and abstract clausal complementation.
Available at http://semlab5.sbs.sunysb.edu/~rlarson/larson-papers.html

von Fintel, Kai. (2001). *Some choicepoints in the syntax/semantics architecture for tense (and aspect)*. Class notes, MIT.

Fox, Danny & Irene Heim. (2006). Class notes for Advanced Semantics, MIT.

Heim, Irene. (1982). *The Semantics of definite and indefinite noun phrases*. PhD dissertation, University of Massachusetts, Amherst.

Heim, Irene. (1999). *Notes on Intensional Semantics*. Class notes for advanced semantics, MIT.

Heim, Irene. (2002). Features of pronouns in semantics and morphology. Talk at USC.

Heim, Irene, & Angelika Kratzer. (1998). *Semantics in generative grammar*. Malden, MA: Blackwell.

Huang, Cheng-Teh James. (1982). *Logical relations in Chinese and the theory of grammar*. PhD dissertation, MIT.

Kamp, J. A. W. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Formal methods in the study of language*, pp. 271-321. Amsterdam: Mathematical Centre.

Kayne, Richard S. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda. (2005). Remarks and Replies: Korean (and Japanese) morphology from a syntactic perspective.

Kratzer, Angelika. (2006). Situations in Natural Language Semantics. In *The Stanford encyclopedia of philosophy*. *Linguistic Inquiry*, 36, pp. 601-633.

Kusumoto, Kiyomi. (2005). On the quantification over times in natural language. *Natural Language Semantics*, 13, pp. 317-357.

Mahajan, Anoop Kumar. (1990). *The A/A-bar distinction and movement theory*. PhD dissertation, MIT.

Mahajan, Anoop Kumar. (2003) Word order and (Remnant) VP movement. In Simin Karimi (ed.), *Word order and scrambling*. Malden, MA: Blackwell.

Matushanski, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry*, 37, pp. 69-109.

Musan, Renate. (1995). *On the temporal interpretation of noun phrases*. PhD dissertation, MIT.

Percus, Orin. (2000). Constraints on some other variables in syntax. *Natural Language Semantics*, 8, pp. 173-229.

Pollock, Jean-Yves, (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, 20, pp. 365-424.

Schlenker, Philippe. (2003). A plea for monsters. *Linguistics and Philosophy*, 26, pp. 29-120.

Soames, S. (1986). Incomplete definite descriptions. *Notre Dame Journal of Formal Logic*, 27, pp. 349-75

Stalnaker, Robert C. (1984). Assertion. Syntax and Semantics, 9. Reprinted in Robert C Stalnaker, *Context and content*. Oxford: Oxford University Press.

Stowell, Tim (1993). Syntax of tense, ms., University of California, Los Angeles.

Stowell, Tim (1995a). The phrase structure of Tense. In J. Rooryck and L. Zaring (eds.), *Phrase structure and lexicon*, pp. 277-291. Dordrecht: Kluwer.

Stowell, Tim (1995b). What do the present and past tense mean? In P. Bertinetto, V. Bianchi, J, Higginbotham, and M. Squartini (eds.), *Temporal reference, aspect and actionality. Vol. 1: Semantics and syntactic perspectives*, pp. 381-396. Torino: Rosenberg and Sellier.

Travis, Lisa deMena. (1984). *Parameters and effects of world order variation*. PhD dissertation, MIT.