

Head Movement as QR

Junri Shimada
junri@mit.edu

0. Overview

- Develops a syntactically natural model where non-individual denoting variables are syntactically present and quantifications over non-individual entities are syntactically realized.
- To that end, I propose that the derivation is driven by head movement, which is an instance of QR.
- Percus's (2000) generalization on the binding constraint for non-individual denoting variables is reduced to an independently motivated locality constraint of head movement.

1. Problem

- ◆ Necessity for syntactic presence of non-individual denoting variables

- Kusumoto (2005)

- (1) a. Hilary married a man who (later) became the president of the U.S.
b. [TP PAST [VP Hilary marry [a man who_i [TP PAST [VP e_i became the president]]]]]

- The relative clause should be interpreted relative to the utterance time, and not the marriage time.

- A Priorian tense logic system with a single time index, where times are manipulated in the meta-language, is not adequate to account for this.

- Percus (2000)

- (2) If every semanticist owned a villa in Tuscany, what a joy this world would be.

- On the transparent reading (if every *actual* semanticist...), *every semanticist* should be interpreted with respect to the actual world, and not counterfactual worlds.

- Again, a system with a single world index is not capable of accounting for this. Since the *if* clause is an island, *every semanticist* should not be able to scope out of the *if* clause.

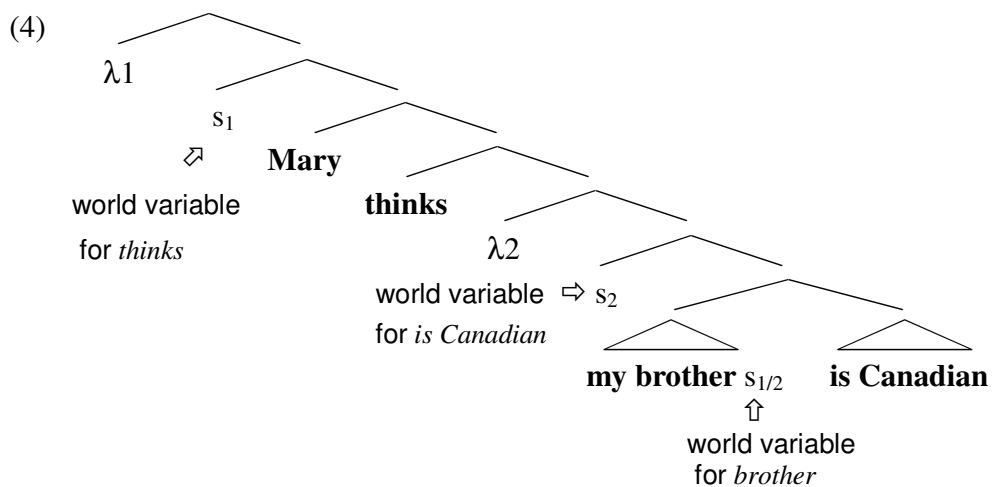
• Simple solution to these data is to assume that the non-individual denoting (time, world, etc.) variables are present in syntax, just like individual denoting variables (such as pronouns).

◆ Percus's problem

• Once we assume that non-individual denoting variables are present in syntax, we need some restrictions on them (Percus 2000).

(3) Mary thinks my brother is Canadian.

• licit configuration



(4) has the meanings in (5a) and (5b) depending on whether the world variable in *my brother* is s_1 or s_2 .

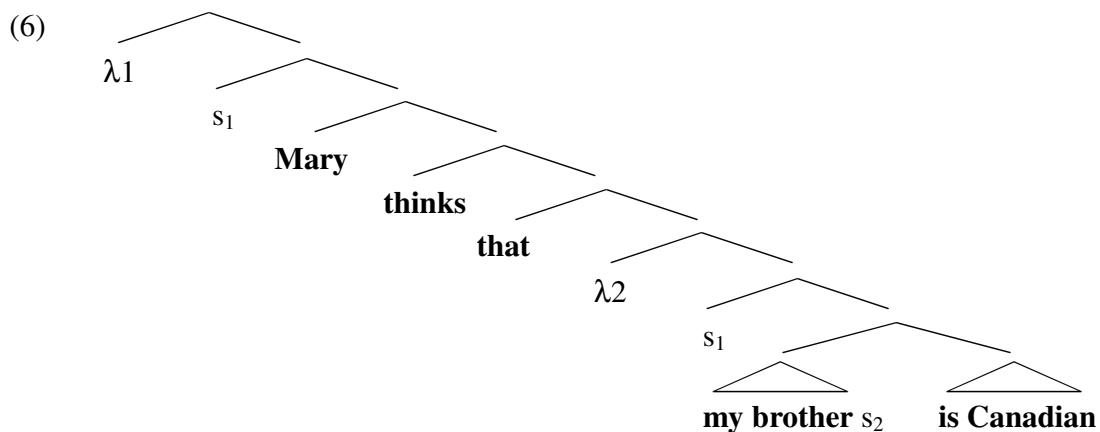
(5) a. For every world w that is consistent with what Mary thinks in the actual world w^* , the unique brother of mine in w^* is Canadian in w .
(*my brother* is transparent)

Scenario: John is my brother in reality. Mary does not know that John is my brother, but she thinks that John is Canadian.

- b. 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 .
(*my brother* is opaque)

Scenario: Mary thinks that John is my brother and that John is Canadian, although John is not my brother in reality.

• illicit configuration



- The world variable for *is Canadian* is not bound by the λ right above it.

- (7) 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^* .

Scenario: John is Canadian in reality. However, in all of Mary's thought worlds, John is not Canadian and he is the unique brother of mine.

- In the scenario provided, the sentence should be true, but it is actually false.

- (8) Percus's 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, but for our purposes here, his "situations" could be regarded as worlds.)

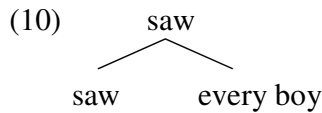
2. Proposal

• If non-individual denoting variables are syntactically present and quantifications over non-individual entities are syntactically realized, it is plausible that the system is similar to what we see with individual denoting variables and quantification over individuals.

- (9) a. John [VP saw [NP every boy]].
 b. LF: [NP every boy]_I [John [VP saw [NP t₁]]].

- *every boy* is generated in the argument position of *saw*, and moves to its scope position at LF due to a semantic type mismatch.

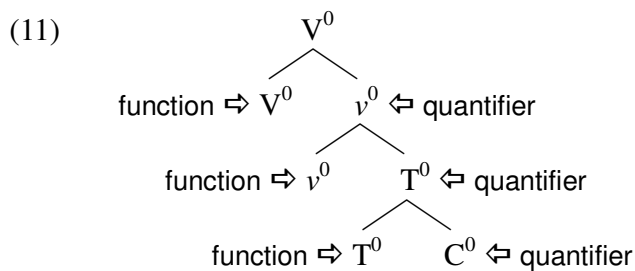
- Since *every boy* is in the argument position of *saw*, their mother is a projection of *saw*.



• In the proposed model, heads that undergo head movement are quantifiers over non-individual denoting entities.

- v^0 : a quantifier over events (type E).
- T^0 : a quantifier over times (type i).
- C^0 : a quantifier over possible worlds (type s).

• The derivation begins with first forming a head complex, by merging heads of the clause in an upside-down and inside-out manner. The whole head complex itself is a head.

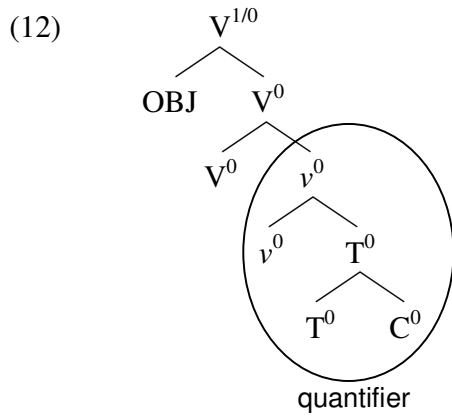


- C^0 is generated in the argument position of T^0 , so T^0 is projected, and so forth.

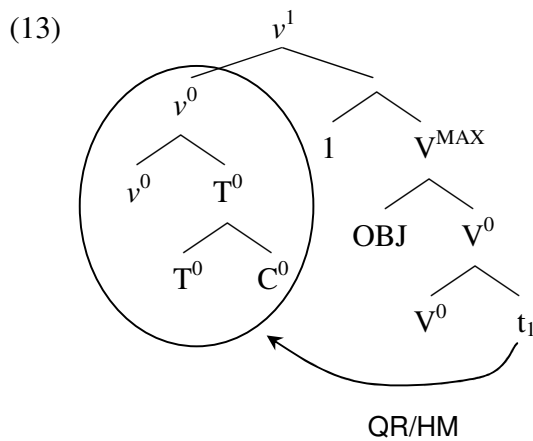
- v^0 , T^0 , C^0 are all forced to undergo QR due to a semantic type mismatch. Head movement is this QR.

◆ The derivation of a simple sentence whose heads are only C, T, v and V.

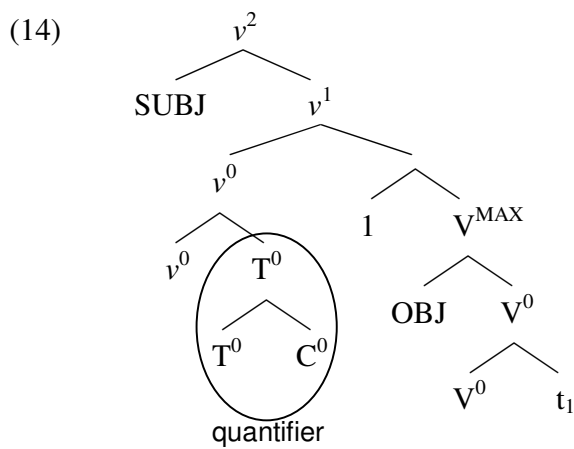
After the complex head in (11) is formed, the object is merged and V gets projected:



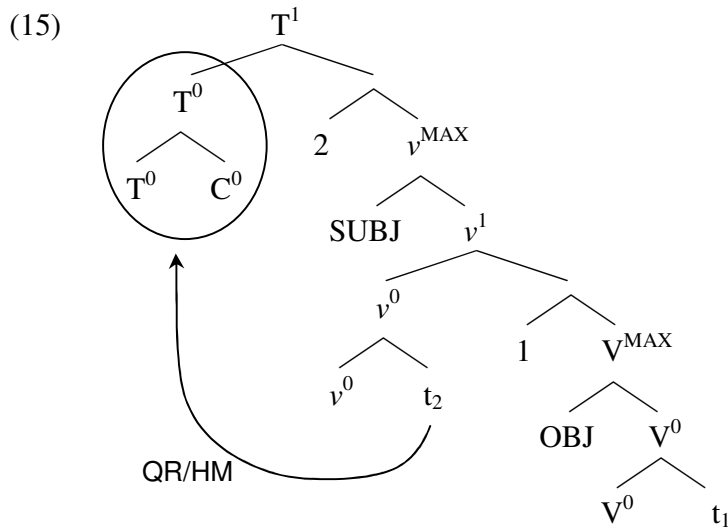
The head complex dominated by v^0 moves by head movement, and v^1 is formed:



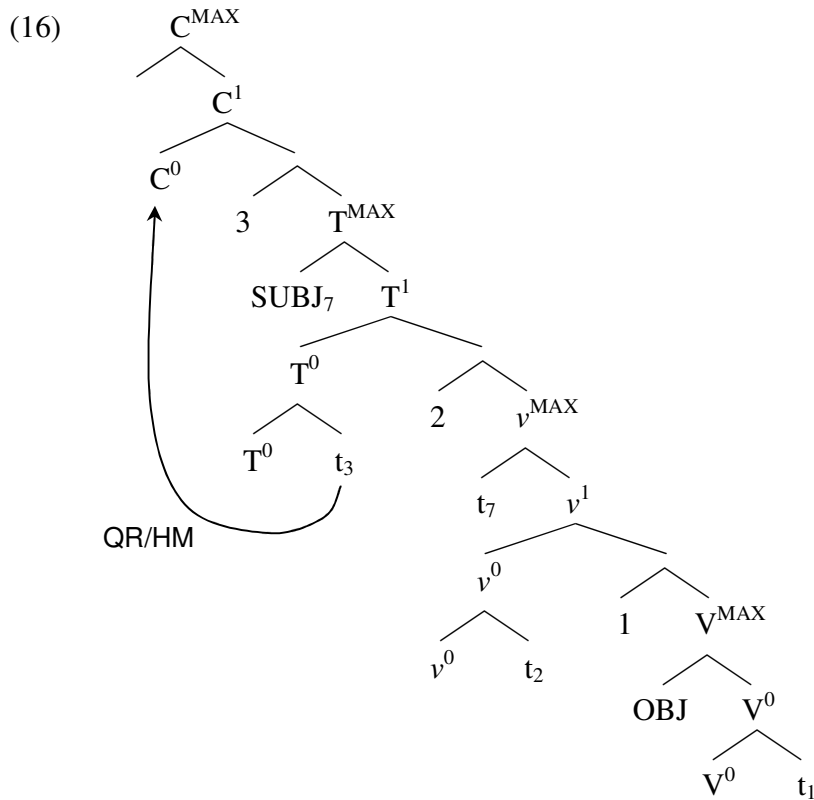
Next, the subject is merged:



The head complex dominated by T^0 moves by head movement:



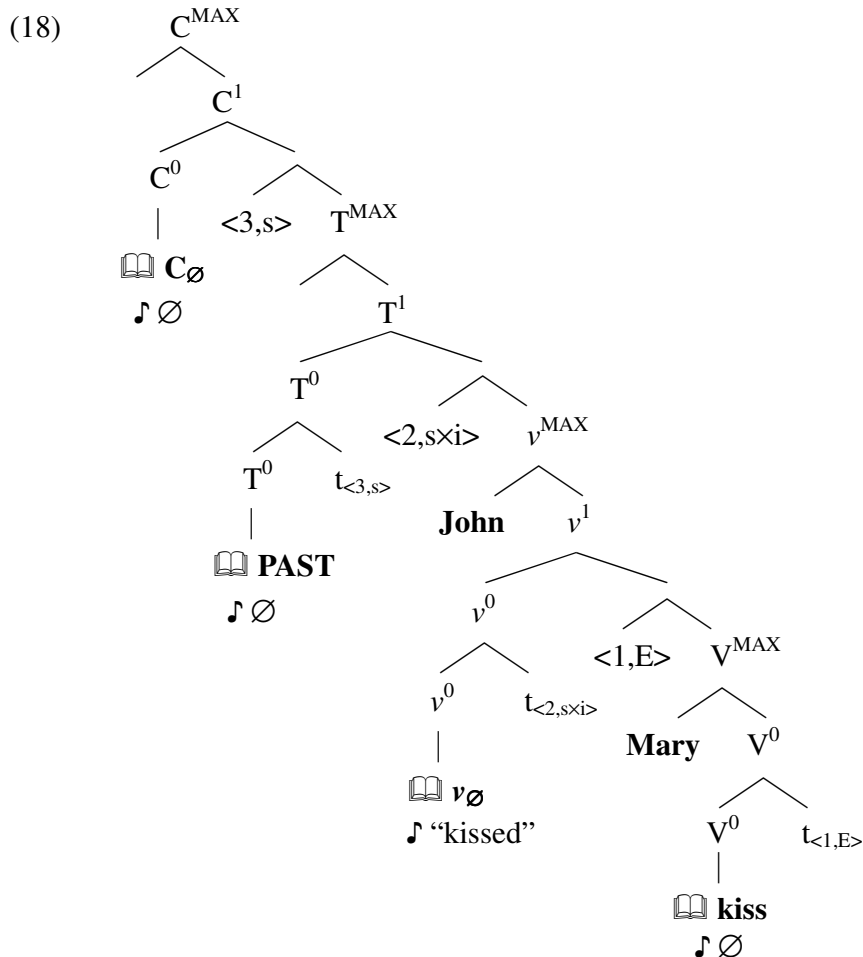
The subject moves to SpecTP by A movement. Then, C^0 moves by head movement:



- The proposed model is compatible with Heim & Kratzer's (1998) system, where variable binding is achieved by virtue of movement.

◆ Semantic composition (simplified version)

(17) John kissed Mary.



(19) $D_{sxi} := D_s \times D_i := \{(x,y) \mid x \in D_s \wedge y \in D_i\}$

(20) Lexical entries

a. $[[\mathbf{PAST}]]^a = \lambda w \in D_s. [\lambda f \in D_{\langle sxi,t \rangle}. \exists t \in D_i [t < \text{NOW} \wedge f((w,t))]]$

b. $[[v_\emptyset]]^a = \lambda \pi \in D_{sxi}. [\lambda f \in D_{Et}. [\lambda y \in D_e. \exists e \in D_E [e \text{ occurs in } \pi \wedge \text{Agent}(e)(y) \wedge f(e)]]]$

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

- T^0 (PAST) takes the world variable as its first argument, pairs it up with the time that it quantifies over, and feeds it to the function f , its second argument. This way, the value of the world variable is passed down in the tree.

- In the denotation of v^0 , π is a pair of a world and a time. v^0 quantifies over events that occur at the time in π and in the world in π (π is understood as a situation).

◆ computation

$$\begin{aligned}
& \llbracket \langle 3, s \rangle \mathbf{T}^{\mathbf{MAX}} \rrbracket^a \\
&= \lambda w \in D_s. \llbracket \mathbf{T}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle}} \\
&= \lambda w \in D_s. \llbracket \mathbf{PAST} \rrbracket^{a^{w/\langle 3, s \rangle}} (\llbracket t_{\langle 3, s \rangle} \rrbracket^{a^{w/\langle 3, s \rangle}}) (\llbracket \langle 2, s \rangle v^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle}}) \\
&= \lambda w \in D_s. [\lambda w' \in D_s. [\lambda f \in D_{\langle s, x_i, t \rangle}. \exists t \in D_i [t < \mathbf{NOW} \wedge f((w', t))]]](w) \\
&\quad (\llbracket \langle 2, s \rangle v^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle}}) \\
&= \lambda w \in D_s. \exists t \in D_i [t < \mathbf{NOW} \wedge \llbracket \langle 2, s \rangle v^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle}}((w, t))] \\
&= \lambda w \in D_s. \exists t \in D_i [t < \mathbf{NOW} \wedge [\lambda \tau \in D_{s, x_i}. \llbracket v^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, \tau/\langle 2, s \rangle}}]((w, t))] \\
&= \lambda w \in D_s. \exists t \in D_i [t < \mathbf{NOW} \wedge \llbracket v^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}}]
\end{aligned}$$

Here,

$$\begin{aligned}
& \llbracket v^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}} \\
&= \llbracket v_{\emptyset} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}} (\llbracket t_{\langle 2, s \rangle} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}}) (\llbracket \langle 1, E \rangle \mathbf{V}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}}) \\
&\quad (\llbracket \mathbf{John} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}}) \\
&= [\lambda \pi \in D_{s, x_i}. [\lambda f \in D_{E, t}. [\lambda y \in D_e. \exists e \in D_E [e \text{ occurs in } \pi \wedge \text{Agent}(e)(y) \wedge f(e)]]]]((w, t)) \\
&\quad (\llbracket \langle 1, E \rangle \mathbf{V}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}})(\mathbf{John}) \\
&= \exists e \in D_E [e \text{ occurs in } (w, t) \wedge \text{Agent}(e)(\mathbf{John}) \wedge \llbracket \langle 1, E \rangle \mathbf{V}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle}}(e)] \\
&= \exists e \in D_E [e \text{ occurs in } (w, t) \wedge \text{Agent}(e)(\mathbf{John}) \\
&\quad \wedge [\lambda e' \in D_E. \llbracket \mathbf{V}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle, e'/\langle 1, E \rangle}}](e)] \\
&= \exists e \in D_E [e \text{ occurs in } (w, t) \wedge \text{Agent}(e)(\mathbf{John}) \wedge \llbracket \mathbf{V}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3, s \rangle, (w, t)/\langle 2, s \rangle, e/\langle 1, E \rangle}}]
\end{aligned}$$

Here,

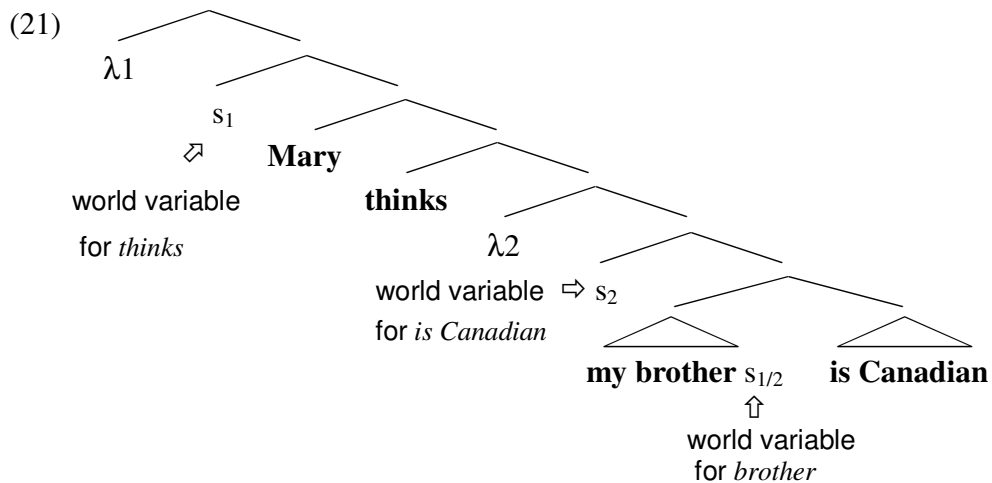
$$\begin{aligned}
 & \llbracket \mathbf{V}^{\mathbf{MAX}} \rrbracket^{a^{w/\langle 3,s \rangle, (w,t)/\langle 2,sxi \rangle, e/\langle 1,E \rangle}} \\
 &= \llbracket \mathbf{kiss} \rrbracket^{a^{w/\langle 3,s \rangle, (w,t)/\langle 2,sxi \rangle, e/\langle 1,E \rangle}} \left(\llbracket t_{\langle 1,E \rangle} \rrbracket^{a^{w/\langle 3,s \rangle, (w,t)/\langle 2,sxi \rangle, e/\langle 1,E \rangle}} \right. \\
 & \quad \left. \left(\llbracket \mathbf{Mary} \rrbracket^{a^{w/\langle 3,s \rangle, (w,t)/\langle 2,sxi \rangle, e/\langle 1,E \rangle}} \right) \right) \\
 &= [\lambda e' \in D_E. [\lambda x \in D_e. \text{Theme}(e')(x) \wedge \text{KISS}(e')]](e)(\text{Mary}) \\
 &= \text{Theme}(e)(\text{Mary}) \wedge \text{KISS}(e)
 \end{aligned}$$

In sum,

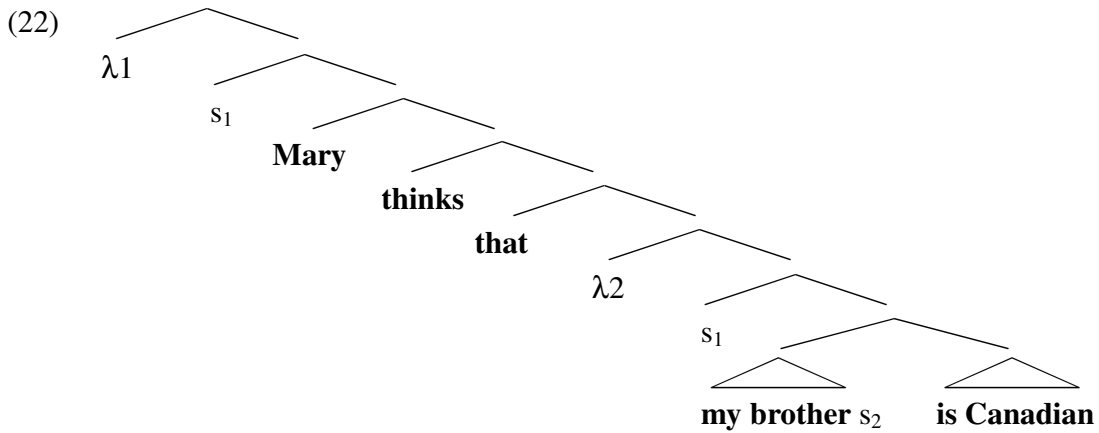
$$\begin{aligned}
 & \llbracket \langle 3,s \rangle \mathbf{T}^{\mathbf{MAX}} \rrbracket^a \\
 &= \lambda w \in D_s. \exists t \in D_i [t < \text{NOW} \wedge \exists e \in D_E [e \text{ occurs in } (w,t) \wedge \text{Agent}(e)(\text{John}) \\
 & \quad \wedge \text{Theme}(e)(\text{Mary}) \wedge \text{KISS}(e)]] \\
 &= \lambda w \in D_s. \text{there is a time } t \text{ such that } t < \text{NOW} \text{ and there is an event } e \text{ such that the agent} \\
 & \quad \text{of } e \text{ is John and the theme of } e \text{ is Mary and } e \text{ is a kissing event that occurs at } t \text{ in } w
 \end{aligned}$$

3. Solution to Percus's problem

- licit configuration



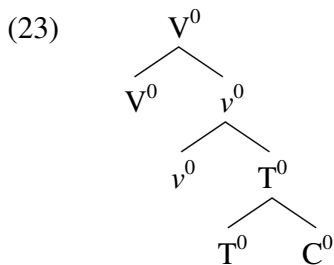
- illicit configuration



- ◆ What accounts for Percus's generalization?

- In my derivation model, the λ operators that bind world variables are introduced by head movement of C^0 .

- Thus, as long as the derivation of a clause begins with the formation of the complex head shown below, Percus's generalization will automatically be observed, since C^0 is in the same head complex as the head of the main predicate (V^0) is.



- Why can't we obtain the illicit binding configuration?

- In (22), the only world variable that $\lambda2$ binds is the world variable in the noun phrase *my brother*.

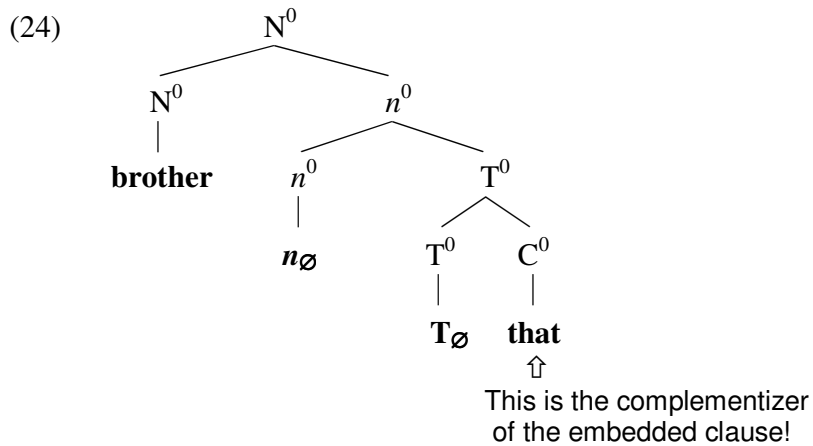
- Since $\lambda2$ is introduced by the head movement of the complementizer of the embedded clause, it must be the case that the complementizer of the embedded clause has moved out of *my brother*.

- This is impossible since *my brother* is a specifier, and head movement is not possible out of a specifier (c.f. Appendix).

- Even if *my brother* itself has moved to some other position, head movement out of *my brother* will never be possible. Since *my brother* is a specifier, it may only move to another specifier.

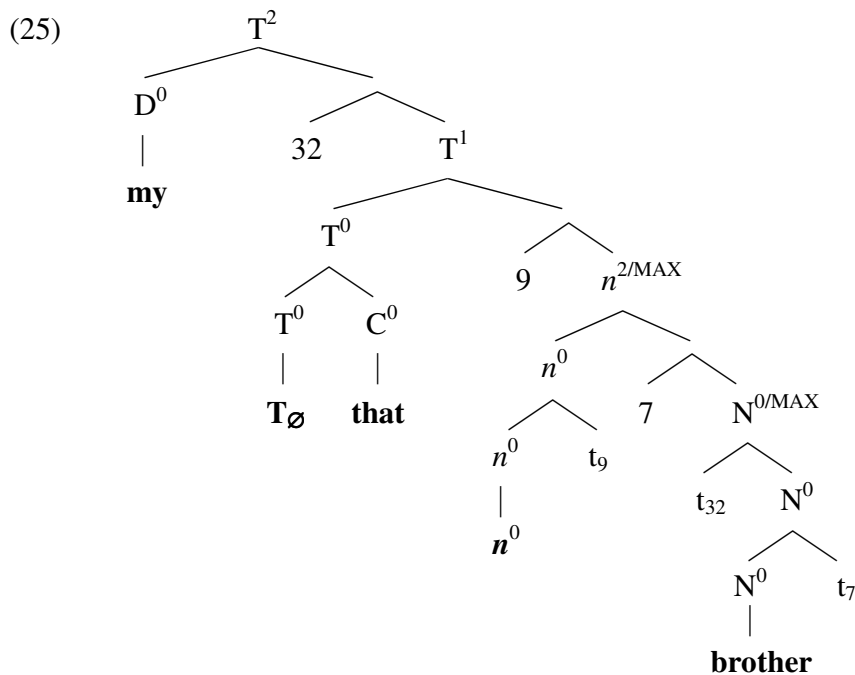
◆ concrete exposition of the (impossible) derivation

The head complex in (24) is created for the subject of the embedded clause:

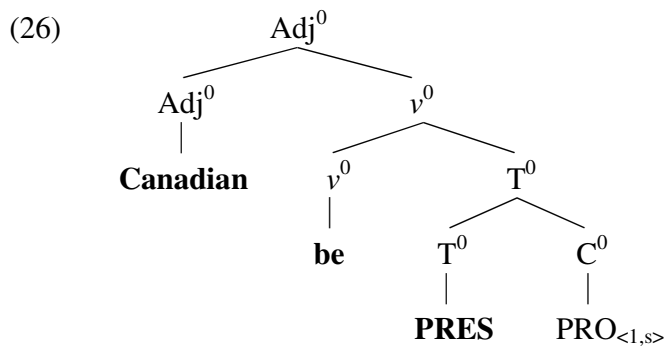


- *n* is to nouns what *v* is to verbs.

The possessive pronoun *my* is a D^0 and is merged with N^0 . After T^0 undergoes head movement, *my* moves to a specifier of T:



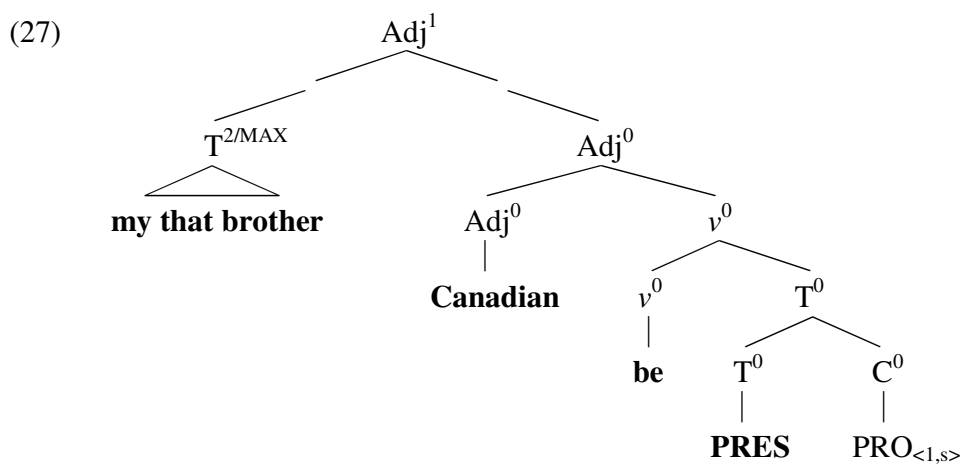
The head complex in (26) is created for the main predicate of the embedded clause:



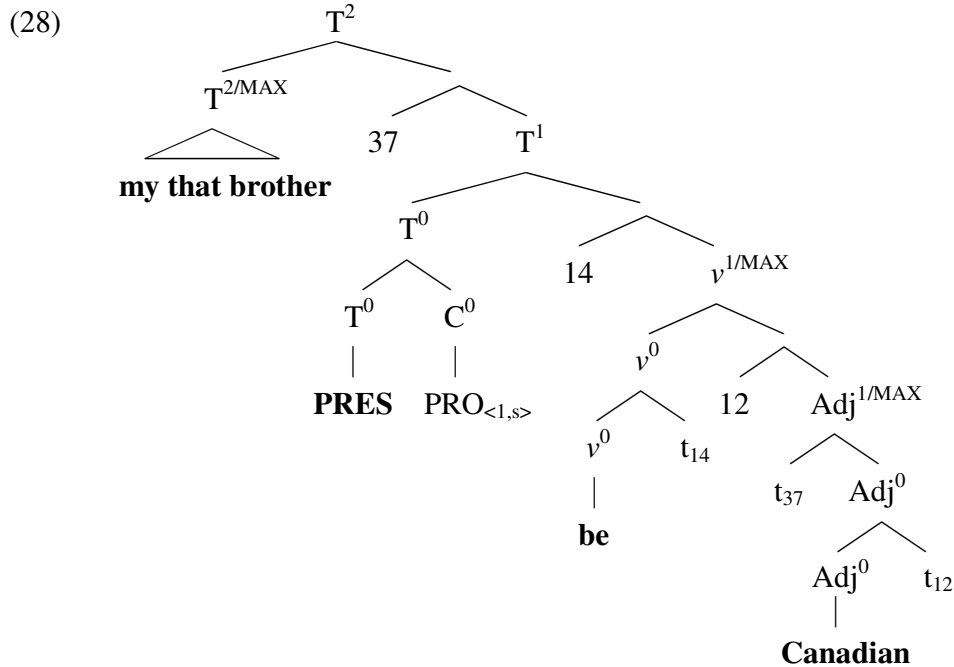
- Since the world variable of the main predicate of the embedded clause is to be bound by $\lambda 1$ in the matrix clause, the complex head for the main predicate of the embedded clause involves a C^0 that is a PRO indexed as 1, which should eventually get bound when head movement of C^0 takes place in the matrix clause.

- I am assuming that the copula *be* is a v^0 .

The subject noun phrase created in (25) is merged with (26) and becomes a specifier:



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:



- Again, the subject noun phrase (T^2) is a specifier.

• Now, to achieve the illicit binding configuration in (22), the C^0 inside the subject noun phrase must undergo head movement and be merged with the root node in (28). However, this is impossible, since the subject noun phrase is a specifier and head movement out of a specifier is impossible.

Concluding remarks

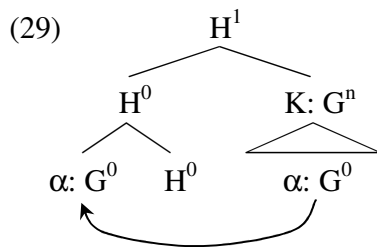
• The proposed model also solves the problem of the standard model of head movement that head movement merges a moving head with a non-root node of a tree, which operation results in a trace that is not c-commanded by the landing site.

• In the proposed model, the derivation begins by forming a head complex and proceeds as heads move out of the head complex by head movement. Heads that undergo head movement are quantifiers over non-individual entities. Thus, head movement is necessary to remedy a semantic type mismatch and it creates binding of non-individual denoting.

• Percus's generalization turns 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 are created by virtue of head movement.

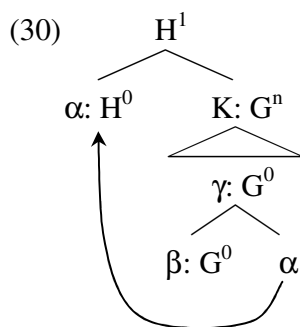
Appendix: locality of head movement (Head Movement Constraint)

- standard model



- What is local with respect to head movement is the head of K.

- my model



- (31) Locality in head movement (this is revised in the paper, but not here)
 Given K, α is local to K with respect to head movement, if there are nodes β and γ such that α and β are the daughters of γ , and β and γ are both heads of K.

References

- 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.
- Heim, Irene, & Angelika Kratzer. (1998). *Semantics in generative grammar*. Malden, MA: Blackwell.
- Kusumoto, Kiyomi. (2005). On the quantification over times in natural language. *Natural Language Semantics*, 13, pp. 317-357.

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.