ECONOMY OF DERIVATION AND ITS APPLICATION

ETSURO SHIMA
Tohoku University

In this paper, I will provide a unified account for the following three seemingly distinct phenomena in terms of Chomsky's (1995) condition on competing derivations and Collins's (1994) Node Traversing Economy: first, the possible order of topics, preposed negatives, and wh-words in pre-sentence positions; second, the proper binding effects in upward movements; finally, the generalization that a CP boundary is a barrier for NP-movement.*

1. Introduction

In this paper, I will provide a unified account for the following three seemingly distinct phenomena: the order of topics, preposed negatives, and wh-words, the proper binding effects in upward movements, and the locality of NP-movements. These phenomena are illustrated by (1)-(4):

(1) a. In the crowd, not a single woman could I find.
   (Langendoen (1979: 432))
   b. *Not a single woman could, in the crowd, I find.

(2) a. These prices, what can anyone do about?
   (Langendoen (1979: 429))
   b. *What can, these prices, anyone do about?

(3) a. ??Which girl₁ do you wonder [which picture of t₁]₂ John likes t₂?
   (Saito (1989: 187))
   b. *[Which picture of t₁]₂ do you wonder which girl₁ John likes t₂?

* This article is a slightly revised version of the paper read at the 12th Annual Meeting of the English Linguistics Society of Japan. Among the many people who have provided me with useful comments, I am most grateful to Professor Masaru Nakamura, Professor Yoshiaki Kaneko, Professor Akira Kikuchi, Sonoko Chiba, and two anonymous EL reviewers. Thanks also go to Peter Robinson for suggesting stylistic improvements. All remaining errors and inadequacies are, of course, my own.
(4) a. It is asked \([\text{CP} \ [\text{how likely } t_1 \text{ to win}]_2 \ \text{John}_1 \text{ is } t_2]\).
b. *\text{John}_1 \text{ is asked } [\text{CP} \ [\text{how likely } t_1 \text{ to win}]_2 \ \text{it is } t_2].

(1) and (2) show that topics must precede both preposed negatives and \(wh\)-words. (3) demonstrates that traces must be properly bound by their antecedents. (4) indicates that NP-movement cannot move across a CP boundary. So far these three phenomena have been given different analyses. However, they have a common property: although each pair of sentences consists of the same lexical items, only sentence (a) is legitimate. Then, I will claim that these three phenomena follow from Chomsky's (1995) condition on competing derivations and Collins's (1994) Node Traversing Economy.

The article is organized as follows: in section 2 I will discuss the possible order of topics, preposed negatives, and \(wh\)-words. In section 3 I will deal with unbound traces, and in section 4 I will discuss the locality of NP-movement. The final section is a summary.

2. The Order of Topic, Neg, and \(Wh\)

Chomsky (1995: 432) suggests the following condition on competing derivations:¹

\[(5) \quad \text{The Condition on Competing Derivations}\]

In evaluating derivations for economy, the competing derivations are determined step-wise; at a particular point in the derivation, we consider the continuations that are permitted, given the initial numeration.

**The Numeration**

Let us take a numeration to be a set of pairs \((l, n)\), where \(l\) is an item of the lexicon and \(n\) is its index, understood to be the number of times that \(l\) is selected.

(Chomsky (1995: 393))

This condition on competing derivations (hereafter the Numeration) guarantees that derivations, which share the same lexical choices, can be compared even if they yield different LF representations.

As the relevant measure for comparing derivations, I adopt the Node Traversing Economy (NTE) proposed by Collins (1994):

¹ In Shima (1993), I independently claim that even if derivations lead to different LF-representations, they can be competing derivations only when they consist of the same lexical items.
(6) **The Node Traversing Economy**
A derivation is optimal if there is no shorter derivation.²
(Collins (1994: 46))

**Length of Derivation**
Derivation D₁ is longer than derivation D₂ if D₁ traverses
more nodes than D₂.³ (Collins (1994: 56))

**Nodes Traversed**
Let D be a derivation, and |Lᵢ|ₜ its links.
Let πᵢ be the path associated with Lᵢ, and Nᵢ be the
cardinality of πᵢ.
The number of nodes traversed is defined as follows:

\[ N_D = \text{the sum of } N_i \text{ for } L_i \text{ in } |L_i|_D \] (ibid.)

**Path**
Let P₁ and P₂ be two categories in a tree.
Let S₁ be the set of categories dominating P₁ and let S₂
be the set of categories dominating P₂.
The path between P₁ and P₂ is defined as follows:

\[ \text{Path} (P₁, P₂) = (S₁ \cup S₂) − (S₁ \cap S₂) \] (ibid.)

Based upon the Numeration and the NTE, let us consider why topics
must precede both preposed negatives and wh-words in pre-sentence
positions. The relevant examples are shown in (1) and (2), repeated
here as (7) and (8) respectively:

(7) a. In the crowd, not a single woman could I find.
b. *Not a single woman could, in the crowd, I find.

² Collins’s (1994: 46) exact formulation of the Node Traversing Economy is as
follows:

(i) **The Node Traversing Economy**
A derivation is optimal if there is no shorter derivation yielding the
same legitimate objects.
(i) means that the condition on competing derivations is to have the same LF-
representation. On the other hand, I assume that competing derivations are
decided by the Numeration. Accordingly, I revise (i) as (6).

³ This formulation for length of derivation is different from the one in Collins’s
(1994: 56)

(i) **Length of derivation**
Derivation D₁ is longer than derivation D₂ if
a. D₁ involves more operations (e.g., Form Chain) than D₂ or
b. D₁ traverses more nodes than D₂.
The difference between (i) and (6) is not relevant to the following argument.
Then, I assume (6) for concreteness.
(8)  a. These prices, what can anyone do about?
   b. *What can, these prices, anyone do about?

The principle of Greed requires every movement to be triggered by the need to check off some feature under SPEC-HEAD agreement. Since Topicalization, Negative Preposing and WH-movement all allow long-distance movements, the driving force of these movements is in C\(^0\) rather than in I\(^0\). Then, I assume that these operations move elements to the SPEC of CP whose head has an operator-feature. The difference in these operations is that Negative Preposing and WH-movement induce the Subject Aux Inversion (SAI), whereas Topicalization does not.\(^4\) Under my assumption, these three operations have the following structures:

(9)  a. \([CP \text{ Nothing}_1 \ [C' \ [C^0 \langle +OP \rangle + \text{did}_2] \ [\text{IP John t}_2 \text{ buy t}_1]]]\)
   b. \([CP \text{ What}_1 \ [C' \ [C^0 \langle +OP \rangle + \text{did}_2] \ [\text{IP John t}_2 \text{ buy t}_1]]?)\]
   c. \([CP \text{ This book}_1 \ [C' \ [C^0 \langle +OP \rangle ] \ [\text{IP John bought t}_1]]]\)

First, let us look at the order of topic—negative. (7a–b) have the common underlying structure in (10):

(10) \([C' \ C^0 \langle +OP \rangle \ [\text{IP I could find not a single woman in the crowd}]]\)

The next operation is to fill the SPEC of C\(^0\). There are two possibilities: the raising of either not a single woman or in the crowd. These possibilities lead to the following derivations (11) and (12), respectively:

(11)  a. \([CP \text{ not a single woman}_1 \ [C' \ [C^0 \langle +OP \rangle + \text{could}_2] \ [\text{IP I t}_2 \text{ find t}_1 \text{ in the crowd}]]]\)
   b. \([C' \ C^0 \langle +OP \rangle \ [CP \text{ not a single woman}_1 \ [C' \ [C^0 \langle +OP \rangle + \text{could}_2] \ [\text{IP I t}_2 \text{ find t}_1 \text{ in the crowd}]]]\)
   c. \([CP \text{ in the crowd}_3 \ [C' \ C^0 \langle +OP \rangle \ [CP \text{ not a single woman}_1 \ [C' \ [C^0 \langle +OP \rangle + \text{could}_2] \ [\text{IP I t}_2 \text{ find t}_1 \text{ t}_3]]]]]\)

(12)  a. \([CP \text{ in the crowd}_3 \ [C' \ C^0 \langle +OP \rangle \ [\text{IP I could find not a single woman t}_3]]]\)

\(^4\) SAI does not take place when either a negative phrase or a wh-phrase appears in a subject position:

(i)  a. Nobody likes John.
   b. Who bought this book?

Following the Vacuous Movement Hypothesis proposed by Chomsky (1986b: 49), I assume that neither negatives nor wh-words in subject positions move to the SPEC of CP, and that they stay in the SPEC of IP.
b. $[C' \cdot C^0 \langle + OP \rangle \ [CP \ \text{in the crowd}_3 \ [C' \cdot C^0 \langle + OP \rangle \ [IP \ \text{I could find not a single woman}_t_3]]]$

c. $[CP \ \text{not a single woman}_1 \ [C' \ [C^0 \langle + OP \rangle + \text{could}_2 \ [CP \ \text{in the crowd}_3 \ [C' \cdot C^0 \langle + OP \rangle \ [IP \ t_2 \ \text{find}_t_1 t_3]]]]$

These are two derivations which can be compared under the Numeration. According to the NTE, in (11), *not a single woman, could*, and *in the crowd* pass across the nodes $[IP], [IP], \text{and } [CP,IP]$, respectively. Then, the number of nodes traversed by these three phrases is four. In (12), in contrast, the same three elements cross the nodes $[IP,CP], [IP,CP], \text{and } [IP]$, respectively, and the number of nodes traversed is five. Then, the NTE selects (11) over (12) since (11) traverses less nodes. Consequently, topics must precede preposed negatives.

Next, let us turn to the order of topics—*wh*-words in (8), repeated here as (13):

(13) a. These prices, what can anyone do about?  
  b. *What can, these prices, anyone do about? 
(13a) and (13b) have the following derivations (14) and (15), respectively:

(14) a. $[C' \cdot C^0 \langle + OP \rangle \ [IP \ \text{anyone can do what about these prices}]]$  
  b. $[CP \ \text{what}_1 \ [C' \ [C^0 \langle + OP \rangle + \text{can}_2 \ [IP \ \text{anyone}_t_2 \ \text{do}_t_1 \ \text{about these prices}]]]$  
  c. $[CP \ \text{these prices}_3 \ [C' \cdot C^0 \langle + OP \rangle \ [CP \ \text{what}_1 \ [C' \ [C^0 \langle + OP \rangle + \text{can}_2 \ [IP \ \text{anyone}_t_2 \ \text{do}_t_1 \ \text{about}_t_3]]]]]$  

(15) a. $[C' \cdot C^0 \langle + OP \rangle \ [IP \ \text{anyone can do what about these prices}]]$  
  b. $[CP \ \text{these prices}_3 \ [C' \cdot C^0 \langle + OP \rangle \ [IP \ \text{anyone can do what about}_t_3]]]$  
  c. $[CP \ \text{what}_1 \ [C' \ [C^0 \langle + OP \rangle + \text{can}_2 \ [CP \ \text{these prices}_3 \ [C' \cdot C^0 \langle + OP \rangle \ [IP \ \text{anyone}_t_2 \ \text{do}_t_1 \ \text{about}_t_3]]]]]$

Unlike Collins (1994: 56), I assume that nonmaximal projections such as $I'$ and $C'$ are not counted as nodes traversed because these categories are invisible for computation (thanks to an anonymous reviewer for bringing this point to my attention).

The same argument applies to the topic—negative sequence in an embedded clause:

(i) a. Fred said that Susan, not once had they seen.  
   (Imanishi (1986: 21), etc.)  
   b. *Fred said that not once had, Susan, they seen.
Given the Numeration, these are competing derivations. In (14), *what*, *can*, and *these prices* pass over *IP*, *IP*, and *IP,CP*, respectively, and the number of nodes traversed is four. In (15), on the other hand, the same three phrases traverse *IP,CP*, *IP,CP*, and *IP*, respectively. The number of nodes is five. Then, the NTE picks out (14) over (15). It follows that topics must precede *wh*-words.7

To summarize, topic—negative and topic—*wh*-word sequences in pre-sentence positions are accounted for in terms of the Numeration and the NTE: the crucial factor is that I0 in SAI crosses less nodes in these sequences than in negative—topic and *wh*-word—topic ones.8, 9

3. The Proper Binding Effects in Upward Movements

Let us turn to the second phenomenon given in (3), repeated here as (16):

3. The Proper Binding Effects in Upward Movements

Let us turn to the second phenomenon given in (3), repeated here as (16):

7 Note that there is a matrix-embedded asymmetry in the order of topics and *wh*-words: while topics must precede *wh*-words in matrix clauses, topics must follow *wh*-words in embedded clauses:

(i) a. *I wonder this book to whom we should give.  
   b. I wonder to whom this book we should give.  (Pesetsky (1989: 22))

The structures of these sentences are as follows:

(ii) a. I wonder [CP this book1 [C C0<+OP> [CP to whom2 [C C0<+OP> [IP we should give t1 t2]]]].
   b. I wonder [CP to whom2 [C C0<+OP> [CP this book1 [C C0<+OP> [IP we should give t1 t2]]]].

(iia–b) are equal in terms of the number of nodes traversed. In (iia), this book passes through *IP,CP* and to whom *IP*. In (iib), this book traverses *IP* and to whom *IP,CP*. Then, in both derivations, the number of nodes is three.

Suppose that wonder requires that the SPEC of its complement is occupied by a *wh*-phrase. This requirement excludes (iia). Therefore, in embedded clauses, *wh*-phrases must precede topics.

8 Neither negative—*wh*-word nor *wh*-word—negative sequence is permissible, as shown in (ia–b):

(i) a. *At no time who has John taken money from?  (Imanishi (1986: 20))
   b. *Who has at no time John taken money from?

(ia–b) have the following structures:

(ii) a. *[CP at no time3 [C [C0<+OP> + [I0+has]1] [CP who2 [C C0<+OP> [IP John t1 taken money from t2 t3]]]]]
   b. *[CP who2 [C [C0<+OP> + [I0+has]1] [CP at no time3 [C C0<+OP> [IP John t1 taken money from t2 t3]]]]]

Neither who in (iia) nor at no time in (iib) is licensed by the complex of C0 and I0. Therefore, these sentences are illegitimate.

9 See Watanabe (1993) for another approach to the order of topic—*wh*-word.
(16) a. Which girl₁ do you wonder [which picture of t₁]₂ John likes t₂?
b. *[Which picture of t₁]₂ do you wonder which girl₁ John likes t₂?

While (16a) is degraded and perhaps has the flavor of a Subjacency violation, (16b) is far worse. Saito (1989) accounts for this fact by resorting to the Proper Binding Condition (PBC) (17), which applies at S-structure:

(17) The Proper Binding Condition:
Traces must be bound. (Fiengo (1977), May (1977))

*Which girl in (16a) binds its trace at S-structure, whereas the one in (16b) does not. Then, the PBC rules out (16b).

Note that Saito’s account presupposes the PBC which is a S-structure condition. Conditions of this kind, however, should be eliminated on the minimalist assumptions that there are no linguistic levels but interface ones, that is, PF/LF levels (Chomsky (1993, 1995)). I will demonstrate that the NTE based upon the Numeration accounts for the contrast in (16a–b), thus eliminating the PBC.¹⁰

(16a–b) share the following structure at a certain stage in their derivations:

(18) [C' C₀ <+OP> [IP John likes [which picture of which girl]]]

There are two possible operations to fill the SPEC of C₀: *which girl* or *which picture of which girl* moves to the SPEC of C₀. The former is illustrated by (19) and the latter by (20):

(19) a. [CP which girl₁ [C' C₀ <+OP> [IP John likes [which picture of t₁]]]]
b. [C' C₀ <+OP> [IP you wonder [CP which girl₁ [C' C₀ <+OP> [IP John likes [which picture of t₁]]]]]]
c. [CP [which picture of t₁]₂ [C' [C₀ <+OP> + I₀₃] [IP you t₃ wonder [CP which girl₁ [C' C₀ <+OP> [IP John likes t₂]]]]]

(20) a. [CP [which picture of which girl]₂ [C' C₀ <+OP> [IP John likes t₂]]]
b. [C' C₀ <+OP> [IP you wonder [CP [which picture of which girl]₂ [C' C₀ <+OP> [IP John likes t₂]]]]]
c. [CP which girl₁ [C' [C₀ <+OP> + I₀₃] [IP you t₃ wonder [CP [which picture of t₁]₂ [C' C₀ <+OP> [IP John likes t₂]]]]]

¹⁰ See Kitahara (1994) and Takano (1994) for different accounts of the PBC effects.
Both derivations have the same three chains: (which girl, t1), ([which picture of t1], t2), and (\textit{t0}3, t3). For the purpose of exposition, we only consider the first two chains. In (19), \textit{which girl} crosses one node [IP] and \textit{which picture of t} three nodes [IP,CP,IP]. In (20), on the other hand, \textit{which girl} traverses two nodes [CP,IP] and \textit{which picture of which girl} one node [IP]. Then, the NTE selects (20) over (19) as economical.\textsuperscript{11}

The PBC effects in upward movements can also be seen in the case of Scrambling:

\begin{enumerate}
\item[(21)] a. *[Mary-ga t2 yonda to]1 sono hon-o2 John-ga t1 Mary-Nom read Comp that book-Acc John-Nom itta. said
\begin{quote}
\textit{\textsuperscript{(Lit.)} [that Mary read t2], that book2, John said t1.}'
\end{quote}
(Saito (1989: 190))

\item[(21)] b. Sono hon-o2 [Mary-ga t2 yonda to]1 John-ga t1 That book-Acc Mary-Nom read Comp John-Nom itta. said
\begin{quote}
\textit{\textsuperscript{(Lit.)} A book1, [that Mary read t2], John said t1.}'
\end{quote}
\end{enumerate}

I assume that scrambling in Japanese is not an optional movement but an obligatory movement like WH-movement, and that a scrambled phrase moves to the SPEC of XP whose head has some feature (cf. Kitahara (1994: 90) and Tomisawa (1994: 287)). Then, (21a-b) share the following underlying structure:

\begin{enumerate}
\item[(22)] \textit{\textsuperscript{[XP sono hon-o2 [X'X0 [IP John-ga [CP Mary-ga sono hon-o yonda to] itta]]]}}
\end{enumerate}

There are two options to fill the SPEC of \textit{X0}, as illustrated in (23) and (24):

\begin{enumerate}
\item[(23)] a. \textit{\textsuperscript{[XP sono hon-o2 [X'X0 [IP John-ga [CP Mary-ga t2 yonda to] itta]]]}}
\item[(23)] b. \textit{\textsuperscript{[XP [CP Mary-ga t2 yonda to]1 [XP sono hon-o2 [X'X0 [IP John-ga t1 itta]]]]}}
\item[(24)] a. \textit{\textsuperscript{[XP [CP Mary-ga sono hon-o yonda to]1 [X'X0 [IP John-ga t1 itta]]]}}
\end{enumerate}

\textsuperscript{11} The selected derivation (20) converges since all features are checked off. But the output of this derivation is deviant because \textit{which girl} moves out of \textit{which picture of which girl} in the SPEC of CP. This extraction violates a kind of Subject Condition. I will not pursue this problem here.
b. \([\text{XP sono hon-o}_2 \text{ [XP [CP Mary-ga t2 yonda to]}_1 \text{ [X X^0 [IP John-ga t1 itta]]]}]\)

In (23), firstly, *sono hon-o* crosses two nodes \(\text{[CP,IP]}\), and then, *Mary-ga t2 yonda to* traverses two nodes \(\text{[IP,XP]}\). In (24), in contrast, the movement of *Mary-ga sono hon-o yonda to* precedes that of *sono hon-o*. The former movement traverses one node \(\text{[IP]}\) and the latter two nodes \(\text{[CP,XP]}\). Then, the NTE favors (24) over (23).  

A similar phenomenon is seen in the case of the interaction of Topicalization and WH-movement, as follows:

(25) *[Which picture of t1]2 do you think that Mary1, John likes t2? 

(25) has the following derivation:

(26) a. \(\text{[C' C^0<+OP> [IP John likes [which picture of Mary]]]}\)

b. \(\text{[CP Mary1 [C' C^0<+OP> [IP John likes [which picture of t1]]]}}\)

c. \(\text{[C' C^0<+OP> [IP you think [CP that [CP Mary1 [C' C^0<+OP> [IP John likes [which picture of t1]]]]]]]}\)

d. \(\text{[CP [which picture of t1]2 [C' [C^0<+OP> + T^0_3] [IP you t3 think [CP that [CP Mary1 [C' C^0<+OP> [IP John likes t2]]]]]]]}\)

In (26a–b), the topic phrase *Mary* moves to the SPEC of embedded C^0, traversing one node \(\text{[IP]}\). In (26c–d), the wh-phrase *which picture of t* passes across four nodes \(\text{[IP,CP,CP,IP]}\) on the way to the SPEC of matrix C^0. Then, the number of nodes traversed is five. Given that the SPEC of C^0 is “visible” to WH-movement, there is another derivation which has the same numeration of (26). Look at derivation (27):

(27) a. \(\text{[C' C^0<+OP> [IP John likes [which picture of Mary]]]}\)

b. \(\text{[CP [which picture of Mary]2 [C' C^0<+OP> [IP John likes t2]]]}\)

c. \(\text{[C' C^0<+OP> [IP you think [CP that [CP [which picture of Mary]2 [C' C^0<+OP> [IP John likes t2]]]]]]}\)

d. \(\text{[CP Mary1 [C' [C^0<+OP> + T^0_3] [IP you t3 think [CP that [CP Mary1 [C' C^0<+OP> [IP John likes t2]]]]]]]}\)

12 One might claim that (24) is excluded by (i):

(i) *Mary-ga sono hon-o yonda to1 John-ga t1 itta.*

This claim is incorrect because (24) does not compete with (i): *sono hon-o* in (i) does not have any feature which is checked off by X^0. Then, (24) and (i) do not have the same numeration.
In (27a-b), the SPEC of embedded C⁰ is occupied by which picture of Mary, which passes through one node [IP]. In (27c-d), the SPEC of matrix C⁰ is filled by Mary, which traverses three nodes [CP,CP,IP]. The number of nodes is four. Therefore, the NTE favors (27) over (26) so that the output of (26), that is, (25) is excluded.¹³

Summarizing, I demonstrated that the PBC effects are accounted for by the Numeration and the NTE. According to the account presented here, the PBC effects can be reduced to the violations of economy conditions. This analysis predicts that when the most economical derivations lead to unbound traces, the PBC effects may be suspended. This prediction is born out. Look at the following sentence:

(28) [How likely t₁ to win]₂ is John₁ t₂?

(Lasnik and Saito (1992: 140), etc.)

Though (28) contains the unbound trace t₁, it is grammatical. This fact is counterevidence to the PBC approach, which requires all traces to be bound by their antecedents at S-structure. My analysis, however, can account for the grammaticality of (28). (28) has the following derivation:

(29) a. [I' [is+I₀] [how likely John to win]]
   b. [IP John₁ [I' [is+I₀] [how likely t₁ to win]]]
   c. [C' C⁰ <+ OP> [IP John₁ [I' [is + I₀] [how likely t₁ to win]]]]
   d. [CP [how likely t₁ to win]₂ [C' [C⁰ <+ OP> + [is+I₀]₃] [IP John₁ [I' t₃ t₂]]]]

Suppose that competing derivations are not only restricted by the Numeration but also by such conditions as Greed and the Extension Condition.¹⁴ Then, there are no alternative derivations to (29). For

¹³ The selected derivation (27) crashes because Mary in the SPEC of the matrix CP cannot check off the feature of the complex of C⁰ and I₀ (see (9)). An anonymous reviewer notes that since the NTE selects the non-convergent derivation (27) as the most economical derivation, my approach implies that economy considerations hold not only of convergent derivations but also of non-convergent ones. I will not go into the problems which arise from my analysis.

¹⁴ Chomsky (1995: 412) suggests that the Extension Condition effects are deducible to more fundamental principles. If Chomsky’s suggestion is correct, competing derivations are restricted by whatever principle accounts for the Extension Condition effects. I leave further investigation of these principles for future research.
example, consider derivations (30) and (31):

(30) a. \([\Gamma \ [is + I^0] \ [how likely John to win]]\)
    b. \([IP \ [how likely John to win]_2 \ [\Gamma \ [is + I^0] \ t_2]]\)
    c. \([C' C^0 \langle + OP \rangle \ [IP \ [how likely John to win]_2 \ [\Gamma \ [is + I^0] \ t_2]]\]
    d. \([CP \ John_1 \ [C' [C^0 \langle + OP \rangle + [is + I^0]_3] \ [IP \ [how likely t_1 to win]_2 \ [\Gamma \ t_3 t_2]]\]

(31) a. \([\Gamma \ [is + I^0] \ [how likely John to win]]\)
    b. \([C' C^0 \langle + OP \rangle \ [\Gamma \ [is + I^0] \ [how likely John to win]]\]
    c. \([CP \ [how likely John to win]_2 \ [C' [C^0 \langle + OP \rangle + [is + I^0]_3] \ [\Gamma \ t_3 t_2]]\]
    d. \([CP \ [how likely t_1 to win]_2 \ [C' [C^0 \langle + OP \rangle + [is + I^0]_3] \ [IP \ John_1 \ [\Gamma \ t_3 t_2]]\]

(30) is blocked by Greed since \(how likely John to win\) moves to the SPEC of IP and \(John\) to the SPEC of CP. (31) is excluded by the Extension Condition, which prohibits the lowering operation of \(John\). Consequently, these two derivations do not compete with (29). Therefore, it follows that (28) is grammatical although it contains the unbound trace.\(^{15}\)

4. The Locality of NP-Movement

Finally, I will demonstrate that the NTE in combination with the Numeration gives an account of Aoun’s (1985) observation that NP-movement cannot cross a CP boundary. This generalization is illustrated by the following sentences:

\(^{15}\) An anonymous reviewer points out that one potential problem with my analysis of (28) has to do with the following facts:

(i) a. *[How likely t_1 to be a riot]_2 is there\_t_2?  
    b. *[How likely t_1 to be taken of John]_2 is advantage\_t_2?  
      (Lasnik and Saito (1992: 141), etc.)

Barss (1986: 436) blocks (i) in terms of the condition on movement in (ii):

(ii) All instances of movement associate two positions X and Y, one of which c-commands the other.  
      (Barss (1986: 408))

Assuming that the NP associated with the expletive \(there\) moves to \(there\) in LF (cf. Chomsky (1986a, 1991)), and that idioms form syntactic units in LF, Barss claims that sentences (i) are illegitimate because the LF-movement of \(a\ riot\) in (ia) and of \(advantage\) in (ib) violates the condition in (ii). It remains to be seen how Barss’s account of (i) interacts with the copy theory of movement which assumes that the operation of Move leaves a copy of the moved element. I leave this problem open.
(32)  a. *John₁ seems [CP that t₁ is sick].
b. *John₁ is illegal [CP t₁ to smoke].
c. *John₁ seems [CP that it is likely t₁ to win].
d. *John₁ is asked [CP [how likely t₁ to win]₂ it is t₂].

The illegitimacy of (32a–b) can be attributed to violations of Greed: John moves from a potential Case position (see Martin (1992), Chomsky (1993, 1995), and Bošković (1993)). As for (32c), John moves to the SPEC of matrix IP, skipping the SPEC of intermediate IP occupied by it, which violates the Minimal Link Condition (MLC). However, neither Greed nor MLC accounts for the ungrammaticality of (32d): in (32d), John neither moves from any potential Case position nor skips any potential landing site. Then, I will propose that the illegitimacy of (32d) is accounted for by the NTE and the Numeration.

(32d) has the following structure at a certain stage of its derivation.

(33)  [I₁' is [AP how likely [IP John to win]]]

The SPEC of IP is filled by either the insertion of it or the raising of John. These two options lead to derivations (34) and (35), respectively.

(34)  a. [IP it is [AP how likely [IP John to win]]]
b. [C₁ C₀<+OP> [IP it is [AP how likely [IP John to win]]]]

16 This argument presupposes that John moves directly to the SPEC of matrix IP. Instead of moving directly to the SPEC of matrix IP, suppose that John first moves to the SPEC of the intermediate clause and then moves to the matrix subject position, as shown below:

(i)  [IP John₁ is likely [CP t₁' that [IP it seems [IP t₁ to be smart]]]]

Since in (i), the first movement from t₁ to t₁' is A'-movement, the intervening Aspecifier it does not interfere with this movement. The next movement from t₁' to John₁ is legitimate A-movement since it does not skip any intervening A-specifier. Then, if (32c) has this derivation, MLC cannot rule out (32c). However, derivation (i) leads to the illegitimate chain (John₁, t₁', t₁) because this chain is not uniform in regard of A/A'-positions: John₁ and t₁ are in A-position, whereas t₁' is in A'-position. Then, this derivation is excluded by whatever principle blocks this type of non-uniform chain.

17 One might argue that the illegitimacy of (32d) lies in the fact that John moves out of how likely John to win in the SPEC of CP. However, this argument does not hold because movements out of phrases in the SPEC of CP produce a milder violation than (32d), as discussed above (see (3) and (16a)). Then, the ungrammaticality of (32d) must be attributed to something other than extractions from the SPEC of CP.

18 See Sakai (1994) for another analysis of (32d).
c. \([\text{CP} [\text{AP how likely [IP John to win]]}_2 [\text{C'} C^0 <+ \text{OP}> [\text{IP it is } t_2]]] \)

d. \([r \text{ is asked [CP [AP how likely [IP John to win]]}_2 [\text{C'} C^0 <+ \text{OP}> [\text{IP it is } t_2]]] \)

e. \([\text{IP John}_1 \text{ is asked [CP [AP how likely [IP } t_1 \text{ to win]]}_2 [\text{C'} C^0 <+ \text{OP}> [\text{IP it is } t_2]]] \)

(35) a. \([\text{IP John}_1 \text{ is [AP how likely [IP } t_1 \text{ to win]]} \]

b. \([\text{C'} C^0 <+ \text{OP}> [\text{IP John}_1 \text{ is [AP how likely [IP } t_1 \text{ to win]]}] \]

c. \([\text{CP} [\text{AP how likely [IP } t_1 \text{ to win]]}_2 [\text{C'} C^0 <+ \text{OP}> [\text{IP John}_1 \text{ is } t_2]]] \)

d. \([r \text{ is asked [CP [AP how likely [IP } t_1 \text{ to win]]}_2 [\text{C'} C^0 <+ \text{OP}> [\text{IP John}_1 \text{ is } t_2]]] \)

e. \([\text{IP It is asked [CP [AP how likely [IP } t_1 \text{ to win]]}_2 [\text{C'} C^0 <+ \text{OP}> [\text{IP John}_1 \text{ is } t_2]]] \)

In (34), \(John\) moves to the SPEC of matrix IP, traversing three nodes \(|\text{IP,AP,CP}|\). By contrast, in (35), \(John\) passes over two nodes \(|\text{IP,AP}|\) on the way to the SPEC of embedded IP. Then, the NTE chooses (35) over (34) so that the output of (34), that is, (32d) is blocked.

Note that I assume that when the SPEC of IP can be filled by the application of either Merge or Move, the NTE decides which operation to choose: the operation is selected which leads to the derivation traversing less nodes. This assumption is different from the one in Chomsky (1995): Chomsky (1995: 428) assumes that when either Merge or Move is available, the former choice is preferred by the principle of Procrastinate if it yields a convergent derivation.\(^19\) My assumption can block (32d), whereas Chomsky’s incorrectly selects derivation (34) and predicts that its output (32d) is legitimate.

My analysis can also account for the contrast in (36a–b):

(36) a. \(*John_1 \text{ is asked [CP [how likely } t_1 \text{ to win]]}_2 \text{ it is believed to be } t_2\). \)

b. \(It \text{ is asked [CP [how likely } t_1 \text{ to win]]}_2 \text{ John}_1 \text{ is believed to be } t_2\).

These two sentences have the following derivations (37) and (38), respectively.

\(^{19}\) I am grateful to an anonymous reviewer for bringing this difference to my attention.
(37) a. \[t' \text{ to be } [\text{AP how likely } [\text{IP John to win}]]\]
b. \[[\text{IP it}_3 \text{ to be } [\text{AP how likely } [\text{IP John to win}]]]\]
c. \[[t' \text{ is believed } [\text{IP it}_3 \text{ to be } [\text{AP how likely } [\text{IP John to win}]]]]\]
d. \[[\text{IP it}_3 \text{ is believed } [\text{IP t}_3 \text{ to be } [\text{AP how likely } [\text{IP John to win}]]]]\]
e. \[[C' C^0 <^+ \text{OP}> [\text{IP it}_3 \text{ is believed } [\text{IP t}_3 \text{ to be } [\text{AP how likely } [\text{IP John to win}]]]]]\]
f. \[[\text{CP } [\text{AP how likely } [\text{IP John to win}]]_2 [C' C^0 <^+ \text{OP}> [\text{IP it}_3 \text{ is believed } [\text{IP t}_3 \text{ to be } t_2]]]\]
g. \[[t' \text{ is asked } [\text{CP } [\text{AP how likely } [\text{IP John to win}]]_2 [C' C^0 <^+ \text{OP}> [\text{IP it}_3 \text{ is believed } [\text{IP t}_3 \text{ to be } t_2]]]\]
h. \[[\text{IP John}_1 \text{ is asked } [\text{CP } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]_2 [C' C^0 <^+ \text{OP}> [\text{IP it}_3 \text{ is believed } [\text{IP t}_3 \text{ to be } t_2]]]\]

(38) a. \[t' \text{ to be } [\text{AP how likely } [\text{IP John to win}]]\]
b. \[[\text{IP John}_1 \text{ to be } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]]\]
c. \[[t' \text{ is believed } [\text{IP John}_1 \text{ to be } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]]]\]
d. \[[\text{IP John}_1 \text{ is believed } [\text{IP t}_1' \text{ to be } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]]]\]
e. \[[C' C^0 <^+ \text{OP}> [\text{IP John}_1 \text{ is believed } [\text{IP t}_1' \text{ to be } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]]]]\]
f. \[[\text{CP } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]_2 [C' C^0 <^+ \text{OP}> [\text{IP John}_1 \text{ is believed } [\text{IP t}_1' \text{ to be } t_2]]]\]
g. \[[t' \text{ is asked } [\text{CP } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]_2 [C' C^0 <^+ \text{OP}> [\text{IP John}_1 \text{ is believed } [\text{IP t}_1' \text{ to be } t_2]]]\]
h. \[[\text{IP It is asked } [\text{CP } [\text{AP how likely } [\text{IP t}_1 \text{ to win}]]_2 [C' C^0 <^+ \text{OP}> [\text{IP John}_1 \text{ is believed } [\text{IP t}_1' \text{ to be } t_2]]]]\]

(37) and (38) are in competition with each other. The difference between these two derivations is the number of nodes traversed by both \textit{it} and \textit{John}. First, look at derivation (37). In (37a-b), the Extended Projection Principle (EPP) in the infinitival clause is satisfied by the insertion of \textit{it}. Then, \textit{it} traverses one node \[\text{IP}\] in (37c-d). Furthermore, in (37g-h), \textit{John} moves to the SPEC of matrix \text{IP}, passing over three nodes \[\text{IP,AP,CP}\]. The number of nodes traversed is four. Next, let us turn to derivation (38). In (38a-b), EPP in the infinitival clause is satisfied by the raising of \textit{John}. This movement traverses two nodes \[\text{IP,AP}\]. In (38c-d), \textit{John} crosses one node \[\text{IP}\]. In (38g-h), \textit{it} is inserted into the SPEC of matrix \text{IP}. The number of
nodes is three. Then, the NTE picks out (38) over (37).

Finally, I want to consider an expletive *there*. Look at sentence (39) and its derivation (40):

\[(39)\] There seems to be a man in the room.
\[(40)\]

\[\begin{align*}
&\text{a. } [I' \text{to } [VP \text{ be } [XP \text{ a man in the room}]]] \\
&\text{b. } [IP \text{ there to } [VP \text{ be } [XP \text{ a man in the room}]]] \\
&\text{c. } [I' [VP \text{ seems } [IP \text{ there to } [VP \text{ be } [XP \text{ a man in the room}]]] ]] \\
&\text{d. } [IP \text{ there}_2 [VP \text{ seems } [IP \text{ t}_2 \text{ to } [VP \text{ be } [XP \text{ a man in the room}]]]]]
\end{align*}\]

In (40a-b), *there* is inserted into the SPEC of IP to check off an EPP-feature on the embedded INFL *to*. Given the initial numeration, there is another way to satisfy EPP on the embedded INFL: the raising of *a man*. This possibility is illustrated by the following derivation:

\[(41)\]

\[\begin{align*}
&\text{a. } [I' \text{to } [VP \text{ be } [XP \text{ a man in the room}]]] \\
&\text{b. } [IP \text{ a man}_1 \text{ to } [VP \text{ be } [XP \text{ t}_1 \text{ in the room}]]] \\
&\text{c. } [I' [VP \text{ seems } [IP \text{ a man}_1 \text{ to } [VP \text{ be } [XP \text{ t}_1 \text{ in the room}]]]]] \\
&\text{d. } [IP \text{ there}_2 [VP \text{ seems } [IP \text{ a man}_1 \text{ to } [VP \text{ be } [XP \text{ t}_1 \text{ in the room}]]]]]
\end{align*}\]

In (40), *there* crosses two nodes \{IP,VP\}. In (41), on the other hand, *a man* passes through two nodes \{XP,VP\}. Then, both derivations are equal with regard to the number of nodes traversed.\(^20\) Suppose that when the NTE cannot decide the most economical derivation, the principle of Procrastinate selects it. Then, (40) is prefered to (41): when either Merge or Move is available for filling the SPEC of IP, (40) selects the former while (41) chooses the latter. Therefore, (40) is more economical so that sentence (39) is produced.

To summarize, I have shown that the Numeration in conjunction with the NTE accounts for the generalization that NP-movement cannot move across a CP boundary.\(^21\)

---

\(^20\) Following a reviewer's suggestion, I assume that only overt movements are subject to the NTE (I am indebted to a reviewer for providing me with this suggestion). Then, the LF movement of the associate *a man* to the expletive *there* is free from this condition. It is an open question why LF movement is not subject to the NTE.

\(^21\) There remains a problem with my analysis. Consider the sentences in (i):

\[(\text{i})\]

\[\begin{align*}
&\text{a. } \text{It}_1 \text{ is believed } [IP \text{ t}_1 \text{ to be told } [that S]_2]. \\
&\text{b. } \ast \text{It}_1 \text{ is believed } [IP [that S]_2 \text{ to be told } t_2].
\end{align*}\]
5. Summary

In this article, I have argued that the interaction of Chomsky's (1995) condition on competing derivations and Collins's (1994) Node Traversing Economy provides a unified account for the seemingly three distinct phenomena. First, this analysis determines the possible order of topics, preposed negatives, and wh-words. Second, it captures the proper binding effects in upward movements. Finally, it accounts for the generalization that a CP boundary is a barrier for NP-movement.

REFERENCES


The derivation producing (ib) traverses less nodes than the one leading to (ia). Then, the NTE incorrectly blocks (ia). This problem is solved by assuming that there is no operation which checks off only an EPP-feature. In other words, an EPP-feature must be checked off by the phrase which has a Case-feature. This assumption favors (ia) over (ib).


Department of English Linguistics
Tohoku University
Kawauchi, Aoba-ku, Sendai-shi
Miyagi 980-77