WEAK CROSSOVER AND SCRAMBLING IN ENGLISH

TORU ISHII
Meiji University*

Keywords: weak crossover, scrambling, D-linking

1. Introduction

The phenomenon of weak crossover (WCO), illustrated in (1), has been under intense investigation for a number of years (see, among others, Koopman and Sportiche (1982) and Mahajan (1990)):

(1) *What does his mother love t?*

Among various proposals in the literature, this paper assumes Mahajan's (1990) condition (2) for expository purposes, though it should be pointed out that the discussion to follow holds valid irrespectively of which approach to the WCO effects is adopted:

(2) A pronoun can be interpreted as a bound pronoun only if it is c-commanded by a binder and its variable (if there is one).

(adapted from Mahajan (1990: 23))

It is well known that A-movement remedies the WCO effects:

(3) Who, t' seems to his mother [t to be smart]

There is, however, a hitherto unnoticed cancellation of the WCO effects, which is exemplified by (4) and (5):

(4) a.*What do you think that his teacher scolded t in yesterday's geology class?

* I would like to thank Brian Agbayani, Hiroshi Aoyagi, and two anonymous EL reviewers for their helpful comments on a previous version of this paper. Remaining errors and omissions are, of course, the sole responsibility of the author. This work was supported in part by the Japan Society for the Promotion of Science under grant Scientific Research C2 14510543 and by a grant from the Institute of Humanities at Meiji University.

© 2003 by the English Linguistic Society of Japan
b. *?Who$_i$ does his$_i$ teacher think that Mary scolded $t_i$ in yesterday’s geology class?

(5) a. Which student$_i$ do you think that his$_i$ teacher scolded $t_i$ in yesterday’s geology class?

b. *?Which student$_i$ does his$_i$ teacher think that Mary scolded $t_i$ in yesterday’s geology class?

In (4) and (5), there is an interplay between the WCO effects and D-linking. While the moved wh-phrases in (5) are D-linked in the sense of Pesetsky (1987), those in (4) are not. In (4), the WCO effects emerge irrespectively of whether the bound pronoun his appears in the embedded clause as in (4a) or in the matrix clause as in (4b). In (5), on the other hand, the WCO effects only emerge when the bound pronoun his appears in the matrix clause as in (5b). When the bound pronoun his appears in the embedded clause, where the wh-phrase originates, the WCO effects are canceled unexpectedly, as in (5a). The following examples illustrate the same point:

(6) a. *?Who$_i$ do you think that his$_i$ teacher will present a special prize to $t_i$ this semester?

b. *?Who$_i$ does his$_i$ classmate think that the teacher will present a special prize to $t_i$ this semester?

(7) a. Which student$_i$ do you think that his$_i$ teacher will present a special prize to $t_i$ this semester?

b. *?Which student$_i$ does his$_i$ classmate think that the teacher will present a special prize to $t_i$ this semester?

(8) a. *?Who$_i$ do you think that his$_i$ paycheck was given to $t_i$ last week?

b. *?Who$_i$ does his$_i$ manager think that the paycheck was given to $t_i$ last week?

(9) a. Which employee$_i$ do you think that his$_i$ paycheck was given to $t_i$ last week?

b. *?Which employee$_i$ does his$_i$ manager think that the paycheck was given to $t_i$ last week?

(10) a. *?Who$_i$ do you think that Mary sent his$_i$ book to $t_i$ yesterday?

b. *?Who$_i$ does his$_i$ wife think that Mary sent the book to $t_i$ yesterday?

(11) a. Which author$_i$ do you think that Mary sent his$_i$ book to $t_i$ yesterday?
b.*?Which author\textsubscript{i} does his\textsubscript{i} wife think that Mary sent the book to \textsubscript{i} yesterday?

This paper argues that this interplay between the WCO effects and D-linking can be given a principled account if we assume that a \textit{wh}-phrase in English undergoes scrambling to the vP-edge position on its way to the final landing site.

2. The Phase Impenetrability Condition and Successive Cyclicity

Before turning to an account of the interplay between the WCO effects and D-linking, a few remarks should be made concerning the Phase Impenetrability Condition (PIC) and successive cyclic movement. Chomsky (2000, 2001a, 2001b) assumes the PIC (12) (adapted from Chomsky (2001a: 14)), which ensures that derivations proceed phase by phase, thereby reducing computational burden:¹

(12) In \([ZP \ Z'c \ [HP \ H' \ YP]]\), where HP is a phase and ZP is the next phase, the domain of H is not accessible to operations at ZP, where phases are vP and CP.

Chomsky claims that the PIC (12) follows from the fact that Spell-Out is subject to the general condition on operations (13) (adapted from Chomsky (2001a: 14)):

(13) A phase Ph\textsubscript{1} is interpreted/evaluated at the next phase Ph\textsubscript{2}.

In order to ensure successive cyclic movement under the PIC, Chomsky assumes the following (adapted from Chomsky (2001a: 12)):

(14) The head of a phase, C and ν, may be assigned an EPP-feature.

This provides an “escape hatch” for successive cyclic movement through the edge. Chomsky assumes the probe-goal theory of movement, where three kinds of uninterpretable formal features are involved in overt movement. In overt \textit{wh}-movement to the Spec of an interrogative C, the following uninterpretable formal features are involved; the Q-feature of C, the \textit{wh}-feature of a \textit{wh}-phrase, and the EPP-feature of C. To allow the probe-goal theory of movement to apply to successive cyclic movement without change, Chomsky (2000) assumes that a non-

¹ Unlike Chomsky (2001a, 2001b), the discussion to follow does not assume a distinction between strong and weak phases.
interrogative C and v without undeleted \( \phi \)-features may also be assigned a non-specific periphery feature (P-feature), which is contingent on the assignment of the EPP-feature to the head of a phase in terms of (14).

While essentially following Chomsky's analysis of successive cyclic movement, however, I depart from Chomsky in claiming that only C, but not v, may be assigned an EPP-feature (and a P-feature). Let us look at how successive cyclic movement proceeds under our analysis, taking (15) as an example:

(15) Where did you buy it t?

During its derivation, we construct the vP phase (16):

(16) \([\textit{vP John} [\textit{v} \{4\} [\textit{VP buy it[\{\phi, \text{Case}\} where[Q, \textit{wh-}]}]]]]\]

In (16), Agree of v with it establishes a Case-agreement relation, deleting the \( \phi \)-features of v and the Case feature of it. Although v is not assigned an EPP-feature or a P-feature, where may raise to the vP-edge position, as shown in (17):

(17) \([\textit{vP where[Q, \textit{wh-}]} [\textit{v'} John [\textit{v} [\textit{VP buy it[\{\phi\} t_{where}]]]]]]\]

Although this movement is not triggered by any feature at this vP-phase level, it does not violate an economy condition which bans superfluous steps in a derivation. This is because according to (13), evaluation/interpretation for the vP phase takes place at the next phase, i.e. the CP phase. That evaluation includes whether or not movement of where to the vP-edge position satisfies the economy condition. At the CP-phase level, C is assigned an EPP-feature and a Q-feature. These features trigger movement of where to the Spec of C, which licenses movement of where to the vP-edge position. This ensures successive cyclic movement without assuming that v may be assigned an EPP-feature (and a P-feature). Note also that this analysis predicts that a moved element cannot remain in the vP-edge position, correctly capturing the lack of Object Shift in English. This is because an object can only move to the vP-edge position if it undergoes feature-driven movement at the next phase.\(^2\)

Our analysis of successive cyclic movement should be preferred over Chomsky's analysis on theoretical grounds in that the former does not assume v with a P-feature. Chomsky (2000) assumes that a

\(^2\) Following Chomsky (2001a), I assume that in languages which allow Object Shift, an object moves to a higher position outside vP, which licenses its movement to the vP-edge position.
P-feature is defective, arguing that a P-feature is analogous for the [per-
son] feature of a defective T. There is a case where C is assigned a
Q-feature, a full complement of peripheral features. By analogy with
the T-system, it is reasonable to claim that we also have a defective C
with a P-feature. In contrast to C, v is never assigned a Q-feature,
which is clear from the fact that a wh-phrase can never be stranded in
the vP-edge position. This casts serious doubt on Chomsky’s analysis,
which assumes v with a P-feature.

I argue that movement of a wh-phrase to the vP-edge position is
scrambling in the sense of Fukui (1993) and Saito and Fukui (1998)
that it is not triggered by any formal feature. Furthermore, like clause-
internal scrambling in Hindi and Japanese (see, among others, Mahajan
(1990) and Saito (1992)), I argue that vP-internal scrambling in English
can be either A- or A'-movement depending on the interpretation of the
moved wh-phrase. The next section will show that our analysis of suc-
cessive cyclic movement can account for the interplay between the
WCO effects and D-linking mentioned above.

3. The Interplay Between the WCO Effects and D-linking

As argued by, among others, Cinque (1990), D-linked wh-phrases
count as specific, since they are understood as presupposing the exis-
tence of a known set of specific elements. It has been claimed by,
among others, Diesing (1992) and Chomsky (2001a) that there is a cor-
respondence between a syntactic position and its specific/non-specific
interpretation. Essentially following Chomsky (2001a), this paper as-
sumes (18):

(18) The non-θ vP-edge position is assigned a specific interpret-
tation.

Given that arguments are A-chains, I claim with Chomsky (2001a) that
while the θ-role of an argument is determined by the position where it
is first Merged, surface interpretation, including specificity, is deter-
mined by the position of the head of an A-chain. Note that unlike
Diesing’s analysis, (18) claims that there is no strict one-to-one cor-
respondence between a syntactic position and its interpretation.
According to (18), it is only the non-θ vP-edge position which receives
a non-ambiguous interpretation as specific. The other positions within
vP have an ambiguous interpretation as either specific or non-specific,
which allows a specific element in-situ to be properly interpreted. Let us first consider (4a). During its derivation, we construct the embedded \( vP \) phase (19):

\[
(19) \quad [v_P \ {\text{who}}_{[Q,\ \text{wh-},\ \text{\epsilon\_case}]} \ [v' \ {\text{his teacher}} \ [{v \ {\text{\#}}} [{v_P \ \text{scolded \ t in yesterday's geology class}}\]])}
\]

In (19), Agree of \( v \) with \( \text{who} \) in its original position establishes a Case-agreement relation, deleting the \( \phi \)-features of \( v \) and the Case feature of \( \text{who} \). Then, the non-D-linked \( \text{wh-phrase} \ \text{who} \) scrambles to the \( vP \)-edge position, which can be either A- or A’-movement. In (19), if scrambling of \( \text{who} \) to the \( vP \)-edge position counted as A-movement, forming an A-chain, the surface interpretation of \( \text{who} \), including specificity, would be determined by the \( vP \)-edge position, which is the position of the head of the A-chain. Since this \( vP \)-edge position is not a \( \theta \)-position, (18) requires \( \text{who} \) to be interpreted as specific.\(^3\) Since \( \text{who} \) is inherently non-D-linked and hence non-specific, this results in a deviant expression, though the derivation converges. If scrambling of \( \text{who} \) counts as A’-movement, on the other hand, \( \text{who} \) forms a trivial one-membered A-chain. The surface interpretation of \( \text{who} \), including specificity, is determined by its original position. (18) is irrelevant and \( \text{who} \) in its original position can be freely interpreted as specific or non-specific. \( \text{Who} \) may be assigned a non-specific interpretation, which is compatible with the inherent property of \( \text{who} \). This results in a convergent derivation with a legitimate interpretation. Hence, scrambling of \( \text{who} \) to the \( vP \)-edge position can only count as A’-movement. Since A’-movement induces the WCO effects, (4a) is deviant.

Let us next consider (5a). During its derivation, we construct the embedded \( vP \) phase (20):

\[
(20) \quad [v_P \ {\text{which student}}_{[Q,\ \text{wh-},\ \text{\#}]} \ [v' \ {\text{his teacher}} \ [{v \ [{v_P \ \text{scolded \ t in yesterday's geology class}}\]}\]]}
\]

In (20), the D-linked \( \text{wh-phrase} \ \text{which student} \) scrambles to the \( vP \)-edge position. Unlike in (19), \( vP \)-internal scrambling may count as A-movement in (20), forming an A-chain. The D-linked \( \text{wh-phrase} \ \text{which student} \) is assigned a specific interpretation due to (18), which is compatible with the inherent property of \( \text{which student} \). This derivation con-

---

\(^3\) Note that this interpretation takes place at the next CP-phase level in accordance with (13).
verges with a legitimate interpretation. Since A-movement remedies the WCO effects, (5a) is acceptable.⁴

Let us finally consider (5b) During its derivation, we construct the matrix vP phase (21):

\[
(21) \quad \left[ vP \quad \text{which student} [Q, \text{wh-}, \#] \quad \left[ v \quad \text{his teacher} \quad [v \quad \text{think} \quad [CP \quad t'' \quad [\text{that} \quad [\text{Mary} \quad [vP \quad t' \quad [v \quad t\text{Mary} \quad [v \quad \text{scolded} \quad t]]]]]]]]]]
\]

In (21), which student first moves to the embedded vP-edge position and then to the Spec of the embedded C on its way to the matrix vP-edge position. Although vP-internal scrambling can be either A- or A'-movement, scrambling of which student from the Spec of the embedded C to the matrix vP-edge position, which crosses over the pronoun his, can only count as A'-movement, inducing the WCO effects. If it counted as A-movement, this would result in an improper movement. This is because which student, which originates in an A-position, would first move into the Spec of C, which is necessarily an A'-position, and then into an A-position. Hence, (5b) is deviant. The examples in (6)–(11) can be explained in the same way.

It should be noted that Chomsky’s (2000, 2001a, 2001b) analysis of successive cyclic movement cannot account for the interplay between the WCO effects and D-linking. Under Chomsky’s analysis, when a wh-phrase enters into a Case-agreement relation with v as in (4) and (5), movement of the wh-phrase to the vP-edge position is triggered by the # and EPP-features of v irrespectively of whether the moved wh-phrase is D-linked or not. This movement counts as A-movement given the widely-accepted assumption that movement triggered by the

---

4 This paper assumes a derivational approach to WCO; condition (2) applies at every point of a derivation (cf. Mahajan (1990)); subsequent movement of his teacher to the Spec of T is irrelevant to such a derivational approach. This view is further supported by (i):

(i) a. Who did Mary think that his brother amused t₁ so much at yesterday’s party?
b. Who did Mary think that his teacher irritated t₁ in yesterday’s geology class?

Given the analysis of psych-verbs proposed by Belletti and Rizzi (1988), where the theme argument his brother originates in a position c-commanded by the experiencer argument who, the derivational approach can correctly predict that (ia, b) exhibit no WCO effects. It remains an unsettled question how we can accommodate such a derivational approach under (13), where interpretation takes place phase by phase. I leave this important issue for further research.
\(\emptyset\)-features on a functional head counts as A-movement (see, among others, Chomsky (2000, 2001a, 2001b)). When a \(wh\)-phrase does not enter into a Case-agreement relation with \(v\) as in (6)–(11), movement of the \(wh\)-phrase to the \(vP\)-edge position is triggered by the P- and EPP-features of \(v\). This movement, triggered by the P-feature on a functional head, counts as A’-movement. Then, under Chomsky’s analysis, there is no way of explaining the contrast between D-linked and non-D-linked \(wh\)-phrases with the WCO effects.

4. Conclusion

This paper has first pointed out the hitherto unnoticed interplay between the WCO effects and D-linking/specificity. I have argued that this interplay can be given an account if we assume that a \(wh\)-phrase in English undergoes scrambling to the \(vP\)-edge position on its way to the final landing site. It has been widely assumed that while languages like Japanese, which have a relatively free word order, have scrambling, languages like English do not. Fukui (1993) and Saito and Fukui (1998) argue that scrambling is a non-feature-driven (optional) movement, proposing a parameter which explains why such an optional movement is allowed in languages like Japanese without violating an economy condition. If the proposed analysis is on the right track, it presents evidence for their view of scrambling, arguing that English also allows scrambling under limited circumstances where it does not violate the economy condition.

Before closing this paper, I will briefly point out that the interplay between the WCO effects and D-linking/specificity is also observed with scrambling.\(^5\) Mahajan (1990) claims that in Hindi, -\(ko\) marked or agreeing objects are interpreted as specific whereas non-agreeing objects without -\(ko\) are interpreted as non-specific. As shown below, there is an interplay between the WCO effects and specificity with Hindi scrambling (Mahajan (1990: 39, 42, 101)):

---

\(^5\) I thank an EL reviewer for bringing this issue to my attention.
(22) *?kOn saa laRkaa uskii maaN t_i ghar se nikaal which boy-Mas his mother home from will throw out-Fem
‘Which boy will his mother throw t_i out of the house?’

(23) a. kis-ko t_i raam-ne socaa [ki uskei bahin-ne t_i dekhaa who Ram-Erg thought that his sister-Erg seen be-past
‘Who did Ram think that his sister had seen t_i?’

b. *kis-ko uskii bahin-ne socaa [ki raam-ne t_i dekhaa who his sister-Erg thought that Ram-Erg seen be-past
‘Who did his sister think that Ram had seen t_i?’

In (22), the non-agreeing object without -ko kOn saa laRkaa ‘which boy-Mas,’ being non-specific, undergoes scrambling. The WCO effects emerge even when the bound pronoun uskii ‘his’ is in the clause where the scrambled phrase originates. In (23), the -ko-marked object kis-ko ‘who,’ being specific, undergoes scrambling. As shown in (23a), the WCO effects are canceled when the bound pronoun is within the embedded clause, where the scrambled phrase originates. As shown in (23b), the WCO effects emerge when the bound pronoun is in the matrix clause. Let us next consider scrambling in Japanese. It has been widely assumed that (24) is an instance of the amelioration of the WCO effects:

(24) Dono gakusei-ni [soitsu-no sidokyokan]-ga ayamatta no which student-Dat he-Gen supervisor-Nom apologized Q
‘To which student did his supervisor apologize t_i?’

Hoji (1995) argues, however, that (24) is not an instance of variable binding, but an instance of what he calls Dem(omstrative)-binding. In (24), the relevant dependency is between the two demonstratives, i.e. do in dono ‘which’ and so in soitsu ‘he,’ rather than between the two DPs, i.e. dono gakusei ‘which student’ and soitsu ‘he.’ Assuming Hoji’s analysis, let us consider (25) and (26) (cf. Hoji (1995: 11)):

(25) a.*?[Kanarinokazu-no kigyo]-ni bengoshi-ga [[soko-no a good number-Gen company-Dat attorney-Nom it-Gen kumiai]-ga t_i kujo-o uttaeta to] omotteiru labor union-Nom complaint-Acc filed that think
‘[A great number of companies], the attorney thinks that its labor union filed a complaint against ti.’

b. *[Kanarinokazu-no kigyo]-ni [soko,-no a good number-Gen company-Dat it-Gen kumiai]-ga [bengoshi-ga t\_i kujo-o uttaeta labor union-Nom attorney-Nom complaint-Acc filed to] omotteiru that think

‘[A great number of companies], its labor union thinks that the attorney filed a complaint against ti.’

(26) a. Dono gakusei,-ni Mary-ga [[[pro\_i hitome mita] which student-Dat Mary-Nom glanced at hito]-ga t\_i muchuninaru to] omotta no person-Nom be fascinated by that thought Q ‘Which student did Mary think that the person who glanced at him would be fascinated by ti?’

b. ?Dono gakusei,-ni [[pro\_i hitome mita] hito]-ga which student-Dat glanced at person-Nom [musume-ga t\_i muchuninaru to] omotta no daughter-Nom be fascinated by that thought Q ‘Which student did the person who glanced at him think that the daughter would be fascinated by ti?’

(25) and (26) suggest that Japanese scrambling also exhibits the interplay. Under Hoji’s analysis, *kanarinonazu-no kigyo* ‘a great number of companies’ in (25) and the empty pronoun *pro* in (26) lack a Demonstrative and thus cannot function as a Dem-binder and a Dem-bindee respectively; the relevant dependency in (25) and (26) is between the two DPs. In (25), where the non-specific quantifier phrase *kanarinokazu-no kigyo-ni* ‘a great number of companies-Dat’ undergoes scrambling, the WCO effects emerge. In (26), where the D-linked (specific) *wh*-phrase *dono gakusei-ni* ‘which student-Dat’ undergoes scrambling, the WCO effects only appear when the bound pronoun *pro* appears in the matrix clause, though the contrast between (26a) and (26b) is not so clear as the proposed analysis predicts. I leave an extended discussion of Japanese scrambling facts for future research.
REFERENCES


Faculty of Literature
School of Arts and Letters
Meiji University
1-1 Kandasurugadai, Chiyoda-ku
Tokyo 101–8301
e-mail: tishii@kisc.meiji.ac.jp