Wh-Category Movement and the Legibility Problem of the Human Language Faculty[1]

Koji ARIKAWA

1. Introduction

We will propose the following hypothesis[2].

(1) *The Overt-Wh-Category Movement Hypothesis*

The legibility conditions imposed by FP at the interface of FL and FP require that a wh-phrase occupy the Spec of CP in the overt syntax, unless Economy principles of C_{HL} forces it to be Q-bound.

FL stands for the human faculty of language (a language organ or a cognitive system which preserves information of sound, meaning and structural organization). FL has an "initial state" S_0 that is an expression of the genes (Chomsky 1998). *FP* stands for the human faculties of performance, and *FP* contains at least the following two types, i.e., modes of perceptual organization (sensorimotor systems, or articulatory-perceptual systems (AP-systems)) and propositional attitudes (systems of thought: beliefs, desires, hopes, fears, ..., or conceptual-intentional systems (CI-systems))[3]. The *overt syntax* stands for the stages of

derivation before Spell-Out, in which all operations visibly affect the surface architecture of the sentence[4]. At Spell-Out, phonetic features are sent to the phonetic-interface level (PF-interface: phonetic form interface) as instructions to sensorimotor-side of FP. Uninterpretable formal features are erased (checked off) by Spell-Out, i.e., within the overt syntax[5]. After Spell-Out, phonetic features being sent to PF-interface and uninterpretable formal features being erased, the syntactic derivation is connected to the semantic-interface level (LF-interface: logical form interface), in which operations do not visibly affect the surface architecture of the sentence, the stages of which are called covert syntax[6]. C_{HL} stands for the computational procedure for human language, which in our purpose contains Economy principles that define the operation Move. A wh-phrase is Q-bound by a functional head F bearing a feature [+wh] if the wh-phrase is c-commanded and coindexed with F[7].

The human species is a mutation in which FL became legible to FP, i.e., FL and FP can mutually access and interact[8]. Optimally, FL must involve two types of features: phonetic features and semantic features. Phonetic features should be legible to sensorimotor system (a part of FP which is connected to modes of perceptual organization), and semantic features should be legible to conceptual system (a part of FP which is connected to propositional attitudes). Legibility conditions had been imposed by FP, which is originally external to FL, so that the information of FL is usable by FP. For example, legibility conditions are connected to the surface interpretive conditions such as topic-comment relation and new-old information, which are imposed by the conceptual-side of FP. More particularly, the principle of full interpretation (FI), which states that all features at LF must be legible to FP, is imposed to LF by CI. The Linear Correspondence Axiom (LCA), which maps a hierarchical relation

to a precedence relation, is imposed to PF by AP[9]. Phonetic and semantic features are interpretable features, i.e., directly legible to FP. If that's the end of the story, we can conclude that human language is a perfect solution to the legibility problem, the essential topic of the minimalist program[10] (Chomsky 1998).

(2) *The Legibility Problem*

How good a solution is FL to the legibility conditions that are imposed by FP?

However, FL involves uninterpretable formal features, which are the driving force of displacement property, one of the important characteristics of human language[11]. Uninterpretable features in FL are simply not legible to FP. Do we then have to conclude that human language is an imperfection because FL contains something uninterpretable to FP?

If the hypothesis in (1) is preserved, the answer to the legibility problem is that the solution is optimal. If wh-category displacement takes place to satisfy the legibility conditions, then the uninterpretable feature [wh], which is the driving force of wh-category movement, is in fact required at the interface. The uninterpretable feature [wh] must somehow be eliminated in the process of computation, since every feature must be interpretable at the interface[12]. How is [wh] erased? [wh] in a wh-phrase which is in a thematic position (a position where the semantic role, e.g., agent, patient, goal, etc., of the wh-phrase is interpreted) is erased if the wh-phrase is attracted to a sufficiently local position to C bearing [wh], i.e., the specifier (Spec) of C. If [wh] in C and [wh] in the wh-phrase match, [wh] is erased. Displacement is the result of the FL-side solution satisfying the legibility conditions imposed by FP. All features used in

FL, both interpretable and uninterpretable, are required in order to satisfy the legibility conditions. The operation Move is relevant to the displacement property, and Move is constrained by Economy principles and several parameter choices. There is only one wh-related operation in C$_{HL}$, i.e., all wh-phrases raise to the Spec of CP to satisfy the legibility conditions, unless other principles of Move force them to remain. FL has solved the legibility problem in a simple, elegant, and natural way, as we will see below. The result is rather surprising: the type of perfection that we are observing here is more akin to the one we find in our inorganic world. Human language is more like a snow flake rather than a giraffe's neck[13].

More particularly, if the hypothesis in (1) is on the right track, it follows that there is neither LF-wh movement (Huang 1982) nor wh-operator movement in the overt syntax (Watanabe 1992). The LF-wh-movement hypothesis claims that English-type language is based on overt-wh movement, whereas Chinese-type language is based on LF-wh movement. Japanese has long been classified as belonging to the latter group. The overt wh-operator movement hypothesis claims that there is something moving in the overt syntax in Chinese-type languages, but it is invisible, i.e., a null-operator or a feature is moving. The overt-wh category movement hypothesis in (1) eliminates the distinction of English-type language and Chinese-type language with respect to wh-feature checking. The elimination of such language-type distinction with respect to syntactic movement is desirable if we assume that there is only one possible human language, apart from the lexicon, i.e., the list of exceptions; arbitrary connections between sound and meaning (Chomsky 1995). We argue that human languages are truly identical with respect to wh-movement, i.e., no language-type distinction such as English-type or Chinese／Japanese-type exists in

this domain of C$_{HL}$.

Before we argue for the Overt-Wh-category Movement Hypothesis, let us eliminate two myths. In the next section 2, we eliminate a myth which is related to the language-type myth, i.e., Japanese-type language exhibits free word order effect, whereas English-type language does not. In section 3, we eliminate the third myth, that the sentence-final element is right-adjoined to the matrix TP. We argue for an analysis in which the sentence-final element is stranded in a Spec of some functional category (= the *Stranding Approach*) (Mahajan 1997).

In Section 4, we will argue for the hypothesis in (1) in detail. By doing so, we will eliminate the myth that there are two types of language, i.e., a language with overt wh-movement and a language with covert wh-movement. Given the *Stranding Approach*, which is argued for in the previous section, an explanation of the "anti-free word order" effect in these languages is provided. We propose a hypothesis that human languages do not differ from each other with respect to wh-movement, i.e., there is no parametric difference among languages in such a way that a language l$_i$ has wh-out-of-place and another language l$_j$ has wh-in-situ. The human language has exactly one type of wh-movement, which is constrained by a set of very few general principles. Wh-movement is forced by the legibility conditions that are imposed to FL by FP: the systems of the brain which are external to FL. As stated at the beginning, when FL became visible in the human brain, FP had to impose these legibility conditions on FL at the interface so that FP could access and interact with FL (Chomsky 1998). Wh-movement is just another example of the general feature-checking operation that is constrained by the general principles.

In Section 5, we will argue that the multiple-wh effect in Japanese is accounted for in a simpler and more natural way, given the *Overt-Wh-*

Category Movement Hypothesis. The standard ECP account is dispensed with. More particularly, it is claimed that the *Extension Condition* (= Cyclicity Requirement) alone accounts for the multiple-wh effect. In Section 6, we will argue that the proposed analysis accounts for the argument-adjunct asymmetry with respect to wh-extraction in Japanese in a simpler and more natural way. Section 7 provides a summary and remaining problems.

2 . No Free Word Order in Japanese - Eliminating the myth of free word order effect

Languages such as Japanese and Hindi have long been considered as "free word order" languages in that the word orders are relatively free[14].

(3) a . Taro-ga Ziro-o korosita (SOV)　　(J)
　　　Taro-nom Ziro-acc killed
　　　'Taro killed Ziro'

　　b . Taro-ga korosita Ziro-o (SVO)
　　c . Ziro-o Taro-ga korosita (OSV)
　　d . Ziro-o korosita Taro-ga (OVS)
　　e . korosita Taro-ga Ziro-o (VSO)
　　f . korosita Ziro-o Taro-ga (VOS)

(4) a . Ram-ne kelaa khaayaa (SOV)　　(H)
　　　Ram-erg banana-abs ate
　　　'Ram ate a banana'

b. Ram-ne khaayaa kelaa (SVO)

c. kelaa Ram-ne khaayaa (OSV)

d. kelaa khaayaa Ram-ne (OVS)

e. khaayaa Ram-ne kelaa (VSO)

f. khaayaa kelaa Ram-ne (VOS)　　　(Mahajan 1990; 19-20)

In contrast, a language like English exhibits rigid word orders[15].

(5) a.?* John Bill killed (SOV)

b. 　John killed Bill (SVO)

c. * Bill John killed (OSV)

d. * Bill killed John (OVS)

e.?* Killed John Bill (VSO)

f. * Killed Bill John (VOS)

However, the "free word order" view of Japanese and Hindi is wrong since the word order is restricted when wh-phrases are involved.

(6) a. Sita-ne kis-ko dhyaan-se dekhaa?　　(H)

　　　 Sita-erg who-acc care-with watch-perf

　　　 'Who did Sita watch carefully?'

b. ??? Sita-ne dhyaan-se dekhaa kis-ko?

　　　 Sita-erg care-with watch-perf who-acc

　　　 'Who did Sita watch carefully?'

　　　　　　　　　　　　　(Mahajan1997a; n.8/1997b; n9)

c. Sita-ne khaaii kyaa ciiz?

Sita-ERG eat-pft-*fem what thing(fem)*

'What did Sita eat?'　　　　　　　　(Mahajan1990:21)

In Hindi, an object with a postposition *ko* fails to agree with the verb (neutral agreement), as in (6a). Such an object with neutral agreement cannot appear post-verbally, as in (6b)[16]. However, interestingly, an object which agrees with the verb can appear in the post-verbal position, as in (6c). A similar restriction is found in Japanese.

（7）a. Hanako-ga dare-o tyuuibukaku mitano?　　　（J）

Hanako-nom who-acc carefully watch-pst-Q

'Who did Hanako watch carefully?'

b. Hanako-ga sore-o naze tyuuibukaku mitano?

Hanako-nom it-acc why carefully watch-pst-Q

'Why did Hanako watch it carefully?'

c. Hanako-ga dono hito-o tyuuibukaku mitano?

Hanako-nom which person-acc carefully watch-pst-Q

'Which person did Hanako watch carefully?'

d. *Hanako-ga tyuuibukaku mitano dare-o?

Hanako-nom carefully watch-pst-Q who-acc

'Who did Hanako watch carefully?'

e. *Hanako-ga sore-o tyuuibukaku mitano naze?

Hanako-nom it-acc carefully watch-pst-Q why

'Why did Hanako watch it carefully?'

  f . * Hanako-ga tyuuibukaku mitano dono hito-o?

   Hanako-nom carefully watch-pst-Q which person-acc

   'Which person did Hanako watch carefully?'

In Japanese, all wh-phrases fail to appear post-verbally, as in (7d-f). English has the following recalcitrant examples.

(8) a . Who saw what?

  b . ?* Who left why?

In English, a wh-phrase such as *what* can appear post-verbally in multiple wh-questions, but a wh-phrase like *why* cannot. Notice that the Minimal Link Condition (MLC) cannot explain the contrast in (8): in both cases, *who* is closer to C when Move (wh) applies[17]. The MLC does not account for the unacceptable word order in Japanese and Hindi (6b, 5d-f) either, since each sentence has only one wh-phrase. It has been claimed that the contrast in (8) is accounted for by argument-adjunct asymmetry with respect to Q-binding: only argument-wh phrases can be Q-bound by C bearing [+wh]. Given the lack of LF wh-movement, the example in (8b) is excluded because the adjunct wh *why* fails to have its feature checked. However, such an explanation based on argument-adjunct asymmetry does not hold for Japanese, since there is no argument-adjunct asymmetry (7d-e). It is a curious situation because the same type of argument-adjunct asymmetry does appear in other areas, both in English and in Japanese; e.g., wh-extraction out of islands.

  Before we investigate the nature of the "anti-free-order" effect of the wh-phrase in these languages, let us argue in the following section that the

post-verbal phrase is not rightward-adjoined to the matrix TP, but it is stranded in some Spec (the *Stranding Approach*), i.e., it is in the position which is hierarchically lowest (Kayne 1994, Mahajan 1997a,b).

3. Against the *Rightward Adjunction Approach* - Eliminating the myth that the sentence-final element is right-adjoined to the matrix TP

Facts of variable binding, anaphor binding, and Condition C effects in Hindi evidence that the post-verbal phrase is stranded in a specifier position (Spec) (the *Stranding Approach*; Kayne1994, Mahajan1997a,b). Japanese is identical to Hindi in these respects, supporting the *Stranding Approach*. The theoretical basis of the *Stranding Approach* is the *Linear Correspondence Axiom* (LCA) proposed in Kayne (1994). Let us review the basic idea of LCA.

3.1. LCA

Let us adopt the following definition of the *LCA* and *Command* (Uriagereka 1998; 537)[18].

(9) *Linear Correspondence Axiom*

A category α precedes a category β if and only if

(a) α asymmetrically commands β, or

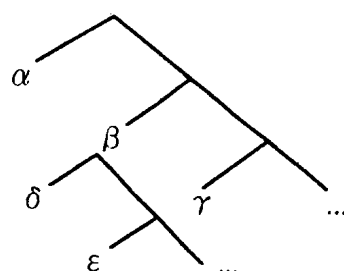(b) γ precedes β and γ dominates α.

(10) *Definition of command*

   Where α and β are accessible to C$_{\text{HL}}$, α *commands* β if and only if

   (a) α does not dominate β, and

   (b) the first category dominating α also dominates β.

The intuitive idea behind the LCA is that the hierarchical structure of a sentence determines the linear ordering of the words in the sentence. In particular, the higher a category appears in a phrase-structure tree, the earlier the category appears in the linear ordering of the sentence. Suppose we have the following schematic structure.

(11)

In (11), α asymmetrically commands β and γ, and β asymmetrically commands γ. β precedes γ and β dominates δ and ε. δ asymmetrically commands ε. The LCA determines that α precedes β and γ, and β precedes γ, and that within β, δ precedes ε. It follows that the linear ordering of terminals α, δ, ε, and γ is $<\alpha, \delta, \varepsilon, \gamma>$. Let us next go over some of the arguments for the *Stranding Approach,* which is based on the LCA.

3.2. Evidence from Variable Binding

Suppose the *Rightward Adjunction Approach* is correct, i.e., a post-verbal phrase is adjoined to the matrix TP[19]. Consider the following

examples.

(12) a . har-ek aadmii-ko$_i$ us$_i$-ke bhaai-ne t$_i$ maaraa　　(H)

　　　　every man-acc his brother-erg hit-perf

　　　　'*His$_i$ brother hit everyone$_i$'

　　 b . * Mohan-ne us$_i$-ke maalik-ko t$_i$ bhejii har-ek kitaab$_i$

　　　　Mohan-erg his owner-dat sent-perf every book

　　　　'Mohan had sent every book$_i$ to its$_i$ owner'

(13) a . daremo$_i$-o soitu$_i$-no hahaoya-ga t$_i$ aisiteiru　　(J)

　　　　everyone-acc his mother-nom love-pres

　　　　'(Lit.) Everyone$_i$, his$_i$ mother loves'

　　 b . * Taro-ga soitu$_i$-no hahaoya-ni t$_i$ hikiwatasita daremo$_i$-o

　　　　Taro-nom his mother-dat introduced everyone-acc

　　　　'(Lit.) Taro introduced to his$_i$ mother, everyone$_i$'

Only an A-element binds a variable[20]. In (12a, 13a), a quantifier phrase (QP) is in the Spec of TP, which is an A-position, rendering variable binding possible. In (12b, 13b), QP is in the right-TP-adjoined position, which is an A'-position, given that no Spec on the right, hence no variable binding[21]. However, the following examples pose a problem.

(14) har-ek kitaab$_1$ [Gita-ne socaa [$_{CP}$ ki Mohan-ne us$_1$-ke maalik-ko t$_1$ bhejii]] (H)

　　　every book-acc Gita-erg thought that Mohan-erg his owner-dat sent

　　　'(Lit.) Every book$_1$, Gita thought that Mohan had given to its$_1$ owner'　　　　　　　　　　　(Mahajan 1997; 100)

(15) daremo₁-o Hanako-ga [CP Taro-ga soitu₁-no hahaoya-ni t₁ hikiwatasita to] itta (J)

everyone-acc Hanako-nom Taro-nom his mother-dat introduced that said

'(Lit.) Everyone₁, Hanako said that Taro had introduced to his₁ mother'

The landing site of long-distance scrambling out of finite clause is necessarily an A'-position, a position from which variable binding is not possible (Mahajan (1990))[22]. The theory then wrongly predicts that (14-15) must be excluded since they involve long-distance scrambling. To explain their acceptable status, we must assume that there is an intermediate A-trace somewhere in a Spec in the embedded clause. Assume that the intermediate A-trace is in the Spec of the embedded-clause TP. This intermediate A-trace would bind the variable. But if such an A-trace is available in the Spec of the embedded TP in (14, 15), so is it in (12b, 13b). The theory again wrongly predicts that the examples in (12b, 13b) should also be acceptable. Thus, by reduction to absurdity, the *Rightward Adjunction Approach* is incorrect[23].

The *Stranding Approach* provides a simple and natural account. In (12b, 13b), the post-verbal QP occupies the Spec of AGRoP (or vP in Japanese), where its Case-feature is checked off. The post-verbal QP is stranded there. The indirect object phrase containing the variable and the verb raise higher. Since a variable must be C-commanded by the antecedent QP (C-command Condition on variables), (12b, 13b) are excluded because they contain an unbound variable. Natural languages do not tolerate such unbound variables. (14, 15) are accounted for if we assume that there is an intermediate trace in the Spec of the embedded TP, which is an A-position[24]. This intermediate A-trace binds the variable. Thus, the

variable binding facts argue for the *Stranding Approach.*

3.3. Evidence from Anaphor Binding

The following contrast indicates that leftward scrambling ameliorates the Binding Condition (A) violation (Mahajan 1997).

(16) a . ??? ek dusure1-ke parivaarō-ne [Sita or Ram]1-ko (khaane ke liye) bulaayaa

　　　　 each other's family-erg Sita and Ram-acc meal-for invited

　　　　 'Each other's family invited Sita and Ram for meal'

　　 b . [Sita or Ram]-ko1 ek dusure1-ke parivaarō-ne t1 (khaane ke liye) bulaayaa

　　　　 Sita and Ram-acc each other's family-erg meal-for invited

　　　　 'Sita and Ram, each other's family invited for meal'

(17) a . ?* otagai1-no kazoku-ga [Taro to Hanako]1-o syokuzi-ni yonda (koto)

　　　　 each other's family-nom Taro and Hanako-acc meal-for invited

　　　　 'Each other's family invited Taro and Hanako for meal'

　　 b . [Taro to Hanako]-o1 otagai1-no kazoku-ga t1 syokuzi-ni yonda (koto)

　　　　 Taro and Hanako-acc each other's family-nom meal-for invited

　　　　 'Taro and Hanako, each other's family invited for meal'

In (16b-17b), the scrambled object (or its trace) must be in the Spec of TP, from which it binds the anaphor. The following examples show that a phrase which is scrambled to an A'-position reconstructs to the original position. Recall that the final landing site of a phrase which has undergone a long-distance scrambling out of a finite clause is necessarily an A'-

Wh-Category Movement and the Legibility Problem of the Human Language Faculty

position (Mahajan 1990).

(18) a . ek dusure1-ke dostõ-ko [Ravi-ne socaa [ki [Ram or Sita]1-ne t̲ bulaayaa thaa]]

    each other's friend-acc Ravi-erg thought that Ram and Sita-erg invited AUX

    'Each other's friends, Ravi thought that Ram and Sita had invited'

b . otagai1-no yuujin-o [Taro-ga [[Ziro to Hanako]1-ga t̲ yonda to] omotta] (koto)

    each other's friend-acc Taro-nom Ziro and Hanako-nom invited that thought

    'Each other's friends, Taro thought that Ziro and Hanako had invited'

In (18), the reciprocal phrase which has undergone a long-distance scrambling reconstructs to the original trace position, thereby being locally bound by the antecedent. Given this much, let us consider post-verbal phrases. Suppose that the post-verbal phrase is rightward-adjoined to the matrix TP, and that there is no Spec on the right. Thus, the rightward-TP-adjoined position is an A'-position. The theory then predicts that the rightward-adjoined post-verbal phrase reconstructs to the original position, where it is locally bound by the antecedent. This prediction is not borne out, however.

(19) a . ??? ek dusure1-ke parivaarõ-ne t̲1 (khaane ke liye) bulaayaa [Sita or Ram]-ko1

    each other's family-erg   meal-for   invited   Sita and Ram-acc

    '(Lit.) Each other's family invited for meal, Sita and Ram'

                                                             (Mahajan 1997)

b . *otagai₁-no kazoku-ga t₁ syokuzi-ni yonda [Taro to Hanako]₁-o

each other's family-nom meal-for invited Taro and Hanako-acc

'(Lit.) Each other's family invited for meal, Taro and Hanako'

In (19), the post-verbal phrase cannot bind the reciprocal. Thus, the *Rightward Adjunction Approach* necessarily leads to a contradiction. By the reduction to absurdity, we must conclude that the *Rightward Adjunction Approach* is wrong.

Let us next suppose that the *Stranding Approach* is correct. The relevant structure of the derivation of (19) is the following.

(20)
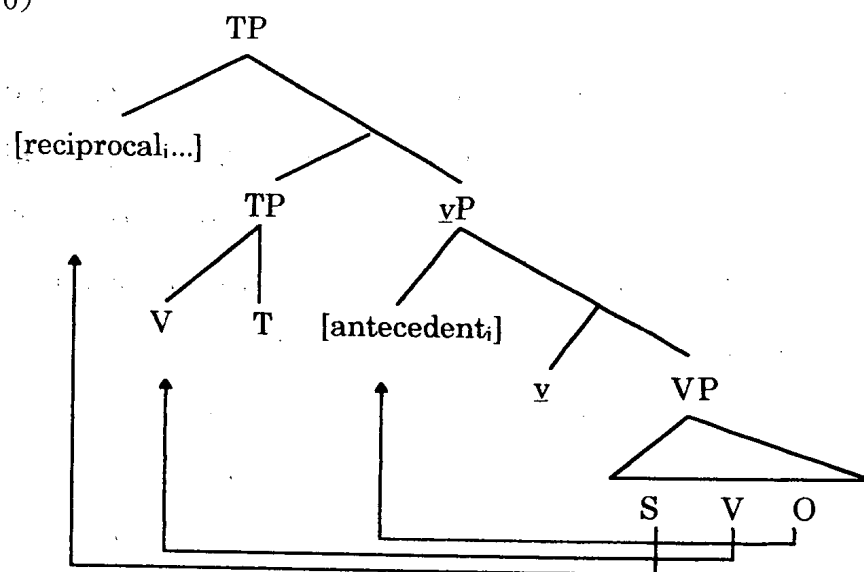
In (20), the reciprocal is not bound by the antecedent, thereby causing the Binding Condition (A) violation. Thus, the *Stranding Approach* does not induce a contradiction.

3.4. Evidence from Condition C Effects

Let us consider how Condition C effects interact with leftward

scrambling. We assume that only an A'-movement shows the reconstruction effect. Let us first observe the following examples of a long-distance scrambling: DO is scrambled out of the finite embedded clause.

(21) a. *us1-ne socaa ki [Gita-ne [tumhaaraa Mohan1-ko likhaa hua p∂tr] phaaR diyaa]

    he-erg thought that Gita-erg your Mohan-dat written letter-acc tore

    'He thought that Gita tore a letter which you had written to Mohan'

   b. *[tumhaaraa Mohan1-ko likhaa hua p∂tr] us1-ne socaa ki [Gita-ne t phaaR diyaa]

    your Mohan-dat written letter-acc he-erg thought that Gita-erg tore

    'A letter which you had written to Mohan, he thought that Gita tore'

(22) a. *[kare1-ga [[Hanako-ga [kimi-ga Taro1-ni kaita tegami]-o yabutta] to] omotta]

    he-nom Hanako-nom you-nom Taro-dat wrote letter-acc tore that thought

    'He thought that Hanako tore a letter which you had written to Taro'

   b. ??*[[kimi-ga Taro1-ni kaita tegami]-o [kare1-ga [[Hanako-ga t yabutta] to] omotta]]

    you-nom Taro-dat wrote letter-acc he-nom Hanako-nom tore that thought

    'A letter which you had written to Taro, he thought that Hanako tore'

Given that a long-distance scrambling is necessarily an A'-movement, the unacceptable status of (21b, 22b) is accounted for if we assume that the scrambled phrase in (21b, 22b) reconstructs to the original trace position, thereby causing a Binding Condition (C) violation: a name is bound. The

following example, however, poses a problem for this analysis.

(23) a. [tumhaaraa Mohan1-ko likhaa hua p∂tr] Gita-ne socaa ki [us1-ne t phaaR diyaa]

　　　 your Mohan-dat written letter-acc Gita-erg thought that he-erg tore

　　　 'A letter which you had written to Mohan, Gita thought that he tore'

　　b. [[kimi-ga Taro1-ni kaita tegami]-o [Hanako-ga [[kare1-ga t yabutta] to] omotta]]

　　　 you-nom Taro-dat wrote letter-acc Hanako-nom he-nom tore that thought

　　　 'A letter which you had written to Taro, Hanako thought that he tore'

Unlike (21b, 22b), the pronoun in (23) is the subject of the embedded clause. If the scrambled phrase reconstructs to the original position in (23), the theory wrongly predicts that the sentences in (23) should be excluded as in (21b, 22b). The unacceptable status of (21b, 22b) and the acceptable status of (23) indicate that the reconstruction site is lower than the matrix subject but at the same time it is higher than the subject of the embedded clause. An intermediate position into which the scrambled phrase drops is the adjoined position of the embedded TP. The following examples indicate that this intermediate position is an A-position.

(24) a. *us1-ne Gita-ko [tumhaaraa Mohan1-ko likhaa hua p∂tr] dikhaayaa

　　　 he-erg Gita-dat your Mohan-dat written letter-acc showed

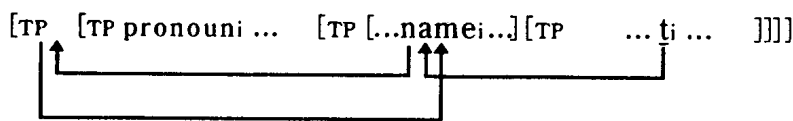　　　 'He showed Gita a letter which you had written to Mohan'

b. [tumhaaraa Mohan1-ko likhaa hua p∂tr] us1-ne Gita-ko t dikhaayaa

   your Mohan-dat written letter-acc   he-erg Gita-dat showed

   'A letter which you had written to Mohan, he showed Gita'

(25) a. *kare1-ga Hanako-ni [kimi-ga Taro1-ni kaita tegami]-o miseta (koto)

   he-nom Hanako-dat you-nom Taro-dat wrote letter-acc showed

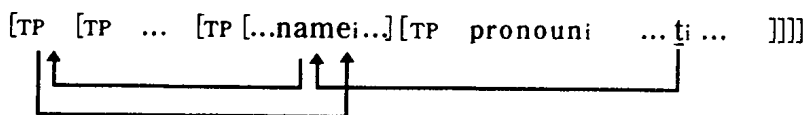   'He showed Hanako a letter which you had written to Taro'

b. [kimi-ga Taro1-ni kaita tegami]-o kare1-ga Hanako-ni t miseta (koto)

   you-nom Taro-dat wrote letter-acc he-nom Hanako-dat showed

   'A letter which you had written to Taro, he showed Hanako'

In (24b, 25b), the phrase which has scrambled to the adjoined position to the local TP can bind the pronoun, indicating that the scrambled phrase is in an A-position. The final LF structures with its derivational history are schematically indicated as follows.

(26) a. (=21b, 22b)

$$[_{TP}\ [_{TP}\text{ pronoun}_i\ ...\ [_{TP}\ [...\text{name}_i...]\ [_{TP}\ \ \ ...\ t_i\ ...\ ]]]]$$

b. (=23)

$$[_{TP}\ [_{TP}\ ...\ [_{TP}\ [...\text{name}_i...]\ [_{TP}\ \text{pronoun}_i\ \ \ ...\ t_i\ ...\ ]]]]$$

The adjoined position to the embedded TP must be an A-position. In (26a), a name is bound, thereby causing the Condition (C) effect. In (26b), a name is free, causing no Condition (C) effect. Now consider the following example.

(27) *Gita-ne use$_i$ t̲ dikhaayaa [tumhaaraa Mohan$_i$-ko likhaa hua p∂tr]

      Gita-erg he-dat showed your Mohan-dat written letter-acc

      'Gita showed to him a letter which you had written to Mohan'

Suppose the *Rightward Adjunction Approach* is correct. Given the preceding discussion, the LF structure of (27) with its derivational history is as follows.

(28)

[$_{TP}$ [$_{TP}$　[...name$_i$...]　[$_{TP}$　...pronoun$_i$ ... t̲ V]]　]
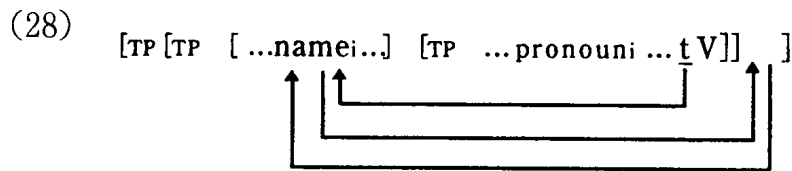
In (28), DO first adjoins to the local TP (A-movement), then it undergoes rightward scrambling (A'-movement), and at LF the post-verbal DO reconstructs to the TP-left-adjoined position. The theory wrongly predicts that the sentence in (27) should be acceptable since the name is unbound, hence no Condition (C) effect. Thus, it is the *Rightward Adjunction Approach* that always leads us to a contradiction. By reduction to absurdity, we conclude that the *Rightward Adjunction Approach* is incorrect. The *Stranding Approach* in contrast provides a simple solution. According to the *Stranding Approach,* the relevant structure of (27) is the following[25].

(29)
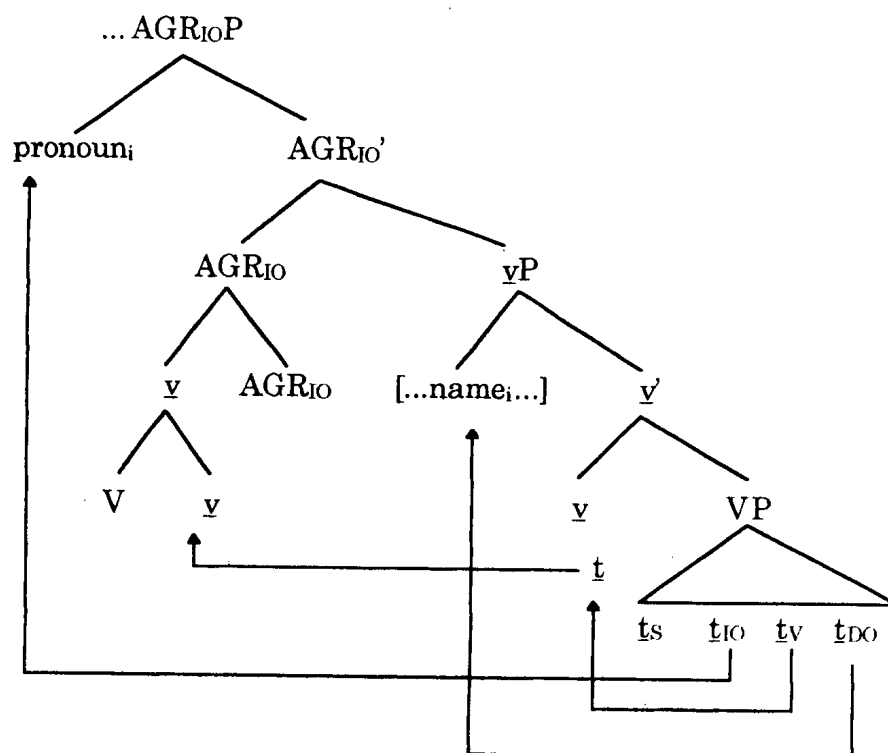
In (29), the name is bound by the pronoun, thereby causing the Condition (C) effect. Hence, the unacceptable status of the example in (27). Thus, the *Stranding Approach* must be correct since it leads us to no contradiction.

The *Rightward Adjunction Approach* necessarily leads the analysis to a contradiction, whereas the *Stranding Approach* provides a natural and simple explanation. We thus adopt the *Stranding Approach* in the following discussion.

4. Wh-in-situ is Moving into the Spec of CP- Eliminating the myth of language-types of wh-movement

4.1. Post-Verbal DO^{wh} is not stranded in the Spec of vP

Let us repeat the relevant examples.

(30)　a. ??? Sita-ne dhyaan-se dekhaa kis-ko?

　　　　Sita-erg care-with watch-perf who-acc

　　　　'Who did Sita watch carefully?' (Mahajan1997a; n.8/1997b;n9)

　　　b. *Hanako-ga tyuuibukaku mitano dare-o?

　　　　Hanako-nom carefully watch-pst-Q who-acc

　　　　'Who did Hanako watch carefully?'

Given the *Stranding Approach,* let us suppose that the post-verbal direct object wh-phrase (DO^{wh}) in (30) raises to the Spec of vP for Case-checking, V-complex (VC) raises to T, subject (S) raises to the Spec of TP, and wh-feature [+WH] is checked by being Q-bound by C bearing [+WH][26]. This yields the order S-V-DO^{wh}, wrongly predicting that the examples in (28b, 29d) are acceptable. The relevant structure is indicated below. The dotted line indicates Q-binding.

(31)
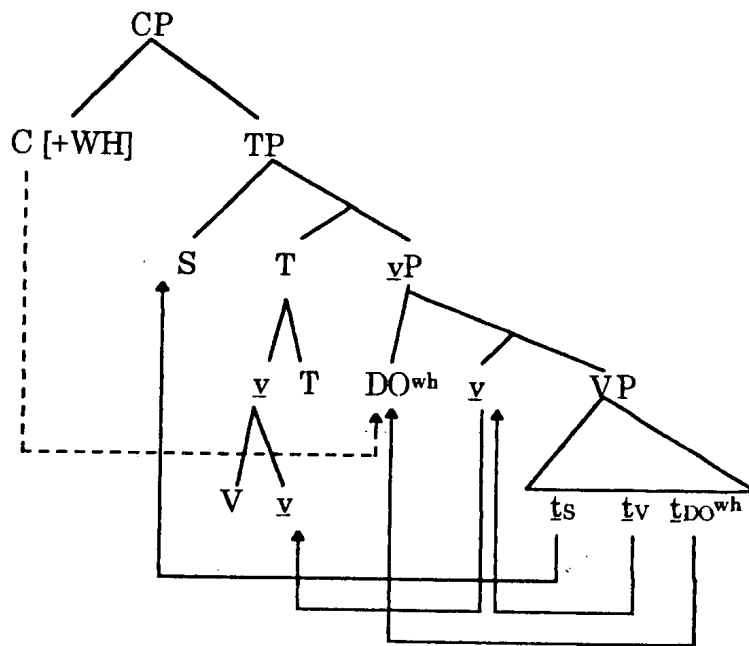
If the structure in (31) were correct (no principle is violated), the theory then wrongly predicts that the sentences in (30) should be acceptable. We conclude that DO^{wh} is not stranded in the Spec of $\underline{v}P$.

4.2. A Proposal

We propose that the post-verbal DO^{wh} in (30) overtly raises to the Spec of CP. The rest follows from the *Linear Correspondence Axiom* (LCA), originally proposed in Kayne (1994). Let us consider the relevant structures of the examples in (28), which are repeated below as (32).

(32) a . ??? Sita-ne dhyaan-se dekhaa kis-ko?

Sita-erg care-with watch-perf who-acc

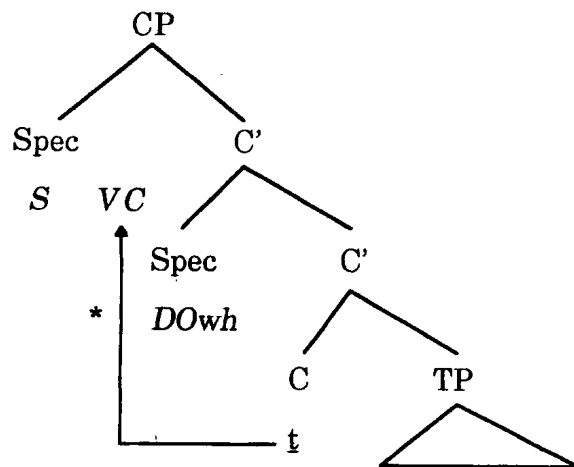'Who did Sita watch carefully?' (Mahajan1997a; n.8/1997b;n9)

b . *Hanako-ga tyuuibukaku mitano dare-o?

Hanako-nom carefully watch-pst-Q who-acc

'Who did Hanako watch carefully?'

We assume that the examples in (32) have the following common structure.

(33)

```
              CP
            /    \
         Spec     C'
                /    \
        S  VC  Spec    C'
            ↑         /   \
        *  DOwh      C     TP
                    /  \
                   |    △
                   t
```

To get VC in front of DO^{wh}, we need a head higher than C. No such head exists, however, given that the Spec precedes head and that C is highest. Raising of VC in (33) is thus superfluous, violating the Economy Principle. Notice that the example in (8b), which is repeated below, is excluded likewise.

(34) ?*Who left why?

In (34), the adjunct wh-phrase *why* is in the Spec of CP. The movement of the verbal complex (VC) is unmotivated since there is no head to which the VC can adjoin, given the universal Spec-Head order. The relevant structure is as follows.

(35)
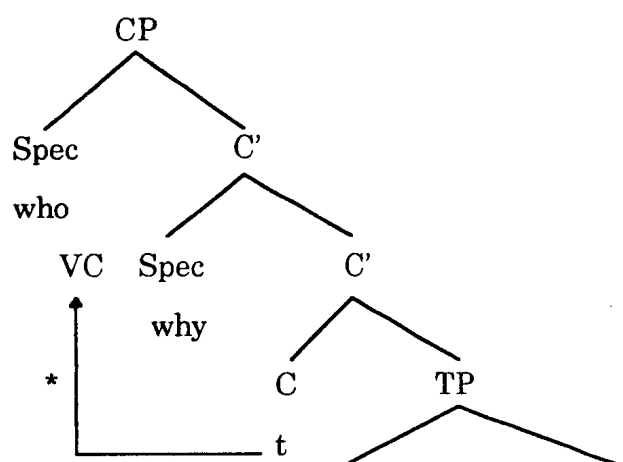
In (35), the verbal complex (VC) $left+\underline{v}+T+C$ is forced to adjoin to a head position between *who* and *why*. The raising cannot be possible since there is no landing site for the VC-movement. If nothing forces it to move, it must stay. If it still moves, the movement is excluded by the Economy Principle.

But why are (36b) and (37b) acceptable?

(36) a . ??? Sita-ne   dhyaan-se dekhaa kis-ko?   (H)

      Sita-erg care-with watch-perf who-acc

      'Who did Sita watch carefully?' (Mahajan1997a; n.8/1997b;n9)

   b . Sita-ne   khaaii kyaa ciiz?

      Sita-erg eat-pft-fem what thing(fem)

      'What did Sita eat?'   (Mahajan1990:21)

(37) a . *John says what? (as a non-echo question)

   b . Who says that?

The difference between (36a, 37a) and (36b, 37b) is that the wh-phrases in the former examples do not agree with the verb, but that those in the

latter do. In (36b), *kyaa ciiz* 'what thing' is in the Spec of AGRoP and is Q-bound there, and in (37b), *who* is in the Spec of TP and is Q-bound there. It appears that a wh-phrase which agrees with the verb is Q-bound. This explains why all wh-phrases raise to the Spec of CP in Japanese; since Japanese lacks agreement-features, no wh-phrase can be Q-bound in this language[27]. The first approximate descriptive generalization is the following.

(38) A wh-phrase which agrees with the verb is Q-bound[28]. Otherwise, a wh-phrase overtly raises to the Spec of CP.

A question comes to mind immediately at this point.

(39) Why is an agreement feature relevant to Q-binding?[29]

Greed, one of the Economy conditions, accounts for (39).

(40) *Greed*

　　Move raises α only if morphological properties of α itself would not otherwise be satisfied in the derivation.　(Chomsky 1995; 261)

Given [+AGR] is a morphological feature, once [+AGR] of a wh-phrase is checked off (erased), there is no reason for the wh-phrase to move further, i.e., it must not move[30]. Thus, the following example is excluded by Economy condition.

(41) * Who$_i$ did t_i say that?

In (41), [+AGR] of *who* is checked off in the Spec of TP, and therefore it must be frozen there. If *who* further raises to the Spec of CP, the movement causes the Economy condition violation. Thus, the example in (41) is excluded along the following examples. *A* in (42b) stands for the "associate" of the expletive.

(42) a. *John seems [(that) t̲ is intelligent]

     b. *there seem [(that) [A a lot of people] are intelligent]

<div align="right">(Chomsky 1995; 261)</div>

In (42a), the morphological properties of *John* are satisfied at the trace position, i.e., [+AGR] of *John* and the lower [Infl-*is*] is checked off. Thus, the raising is barred by Greed. In (42b), the morphological properties of the associate *a lot of people* are satisfied at the subject position of the embedded clause, i.e., [+AGR] of the associate a *lot of people* and the lower [Infl-*are*] is checked off. Thus, the raising is barred by Greed. If the movement is barred, Case and ϕ-features of the matrix [Infl-*seem*] will not be checked off, causing a crash[31]. The following examples support our analysis.

(43) a. *John says what? (a non-echo question)

     b. Who says what?

In (43a), *what* does not agree with the verb, therefore it cannot be Q-bound in the Spec of v̲P (= AGRoP). Since *what* in (43a) remains in the Spec of AGRoP without any reason, it violates the Economy conditions[32]. In (43b), *what* is forced to be Q-bound in the Spec of v̲P since the movement of *what* would cause the Minimal Link Condition violation. If

what overtly raises to the Spec of CP, the sentence is excluded by the MLC. The wh-phrase *what* "looks ahead" to the result of the MLC violation after raising, and it chooses to be Q-bound in a last resort fashion[33]. The relevant structure of (43b) is shown in the following. The dotted lines indicate Q-binding.

(44)
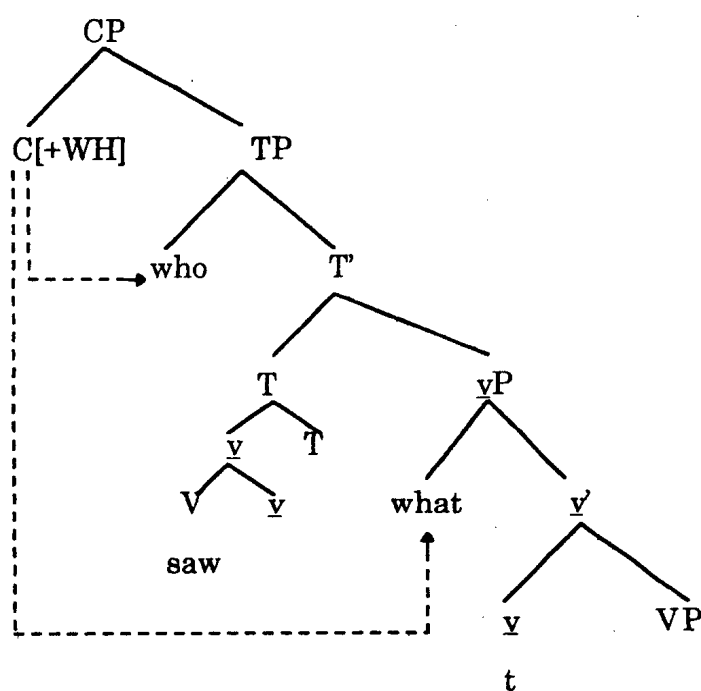
If *what* raises to the Spec of CP, *what* would cross over *who,* which is closer to C[+WH], thereby inducing the MLC violation (* *What did who see?*). Thus, the Last Resort Condition, one of the Economy conditions, forces *what* to remain in the Spec of vP. The following examples in English further support the view that the MLC is relevant.

(45)　a ．*He has said who$_i$ who loved t$_i$

　　　b ．Which book$_i$ did which person buy t$_i$?

　　　c ．He has said which student$_i$ which professor loved t$_i$?

The example in (45a) is excluded by the MLC: a non-closer wh-phrase *who*i is raised in violation of the MLC. In (45b-c), the MLC (distance calculation) is irrelevant, since in a wh-phrase of the *which* x type, the wh *which* no longer C-commands the other wh-phrase (Epstein 1993, Kitahara 1993). *Which person* in (45b) and *which professor* in (45c) are Q-bound in the Spec of TP. *Which book* in (45b) and *which student* in (45c) raise to the Spec of CP passing the wh-phrase in the Spec of TP, given the irrelevance of MLC and the lack of agreement features.

Furthermore, the theory predicts that a phrase without [+wh] can appear post-verbally in Japanese. This is so because a wh-phrase as such need not raise to the Spec of CP, and thus can remain in the Spec of vP. The prediction is borne out.

(46) a. Taro-wa kekkyoku mitano sono nani-o?

       Taro-TOP after all saw Q that doodad

       '(So,) did Taro see that doodad after all (I forgot the name of the object)?'

   b. Taro-wa kekkyoku mitano sono hon-o?

       Taro-TOP after all saw Q that book-acc

       '(So,) did Taro see that book after all?'

Since *sono nani-o* 'that doodad-acc' in (46a) and *sono hon-o* 'that book-acc' in (46b) bear no [+wh], the raising to the Spec of CP is needless, i.e., rather it must not raise to the Spec of CP, given the Economy conditions. They remain in the Spec of vP, where its Case-feature is checked off. The verbal complex (VC) adjoins to T, yielding S-V-O order[34].

If the analysis is on the right track, (38) can be restated as follows.

(47) *Universal WH-Licensing Condition*

Unless the movement violates some Economy conditions,

a wh-phrase α either

(a) raises to the Spec of CP in the overt syntax if α bears [−AGR], or

(b) is Q-bound by the local C with [+Q] if α bears [+AGR].

5 . A Simple FL-Solution: Multiple Wh Effect

In this section, we will show that the *Overt-Wh-Category Movement Hypothesis* predicts that what have long been considered as recalcitrant phenomena are another optimal solutions by FL to the legibility problem. Multiple-wh effect in Japanese, which is shown in (48a-b), is one of them. The contrast in (48b-c) indicates that focus with stress is relevant. Interestingly, multiple-wh contrast disappears in post-verbal environments, as in (48d-e).

(48)　a . Hanako-ga nani-o naze tabetano?

　　　　　Hanako-nom what-acc why eat-pst-Q

　　　　　'What and why did Hanako eat?'

　　　b . * Hanako-ga naze nani-o tabetano?　　(*Naze* unstressed.)

　　　　　Hanako-nom why what-acc eat-pst-Q

　　　c . Hanako-ga *NAZE NANI*-o tabetano?　(*Naze* and *nani* stressed.)

　　　　　　　　why　　what-acc　　-Q

　　　　　　　　　　　　(Saito1982,1992, Watanabe1992)

d . * Hanako-ga naze tabetano nani-o?

       why   -Q what-acc

e . * Hanako-ga nani-o tabetano naze?

       what-acc   -Q why

Watanabe (1992) proposed the Anti-Superiority Condition, as in (49), to explain the multiple-wh effect.

(49) *Anti-Superiority Condition* (ASC)

    A multiple question is well-formed only if there is a wh-phrase which is not c-commanded by the wh-phrase that is moved first.

The *wh-phrase that is moved first* is typically an adjunct-wh phrase, which is forced by the ECP (Empty Category Principle), i.e., an adjunct-wh trace, but not an argument-wh trace, must be antecedent-governed by the antecedent which assigns its index to CP. CP is assigned the index of a wh-phrase that is raised to the Spec of CP first[35]. The ASC states that an argument-wh phrase which c-commands an adjunct-wh phrase saves the sentence from the multiple-wh effect. In (48a), the argument-wh *nani-o* 'what-acc' c-commands the adjunct-wh *naze* 'why.' Thus, (48a) is permitted by the ASC. In (48b), there is no argument-wh phrase which c-commands the adjunct-wh phrase. Thus, (48b) is excluded by the ASC. However, (48c) and (48e) pose a problem for the ASC. The ASC wrongly predicts that (48c) is excluded and that (48e) is permitted, given the *Stranding Approach.* Saito (1992) generalized the LF Saving Effects, as in the following.

(50) *Requirements for the LF saving effects*

β (an adjunct wh phrase within an island) is saved by

α (an argument wh phrase) iff

（ａ）both α and β are subject to Move at LF and

（ｂ）α and β are clausemates at D-structure and

（ｃ）α and β are clausemates at S-structure and

（ｄ）α precedes β at S-structure

The *Requirements for the LF saving effects* account for the contrast in (48a-b) and the example in (48d): the examples in (48b) and (48d) do not satisfy the requirement (50d). However, the *Requirements for the LF saving effects* wrongly predict that (48c) should be excluded and that the example in (48e) should be permitted. Thus, the ASC and the *Requirements for the LF saving effects* must contain an incorrect formulation.

We propose that the *Overt-Wh-Category Movement Hypothesis* and the *Extension Condition* account for (48); no new conditions are necessary. The FL solution is simple: (48d-e) are excluded by the *Economy principle:* they involve unmotivated head movement, as in (6b), (7d-f), and (8b), repeated below.

(51) ａ. ??? Sita-ne dhyaan-se dekhaa kis-ko?　　（=6b）

Sita-erg care-with watch-perf who-acc

'Who did Sita watch carefully?'　(Mahajan1997a; n.8/1997b;n9)

ｂ.＊Hanako-ga tyuuibukaku mitano dare-o?　　（=7d）

Hanako-nom carefully watch-pst-Q who-acc

'Who did Hanako watch carefully?'

c . * Hanako-ga sore-o tyuuibukaku mitano naze?   (=7e)

Hanako-nom it-acc carefully watch-pst-Q why

'Why did Hanako watch it carefully?'

d . * Hanako-ga tyuuibukaku mitano dono hito-o?   (=7f)

Hanako-nom carefully watch-pst-Q which person-acc

'Which person did Hanako watch carefully?'

e . ?* Who left why?   (=8b)

(48d-e) and (51) involve an unmotivated head movement, as in the following.

(52) (=33, 35)
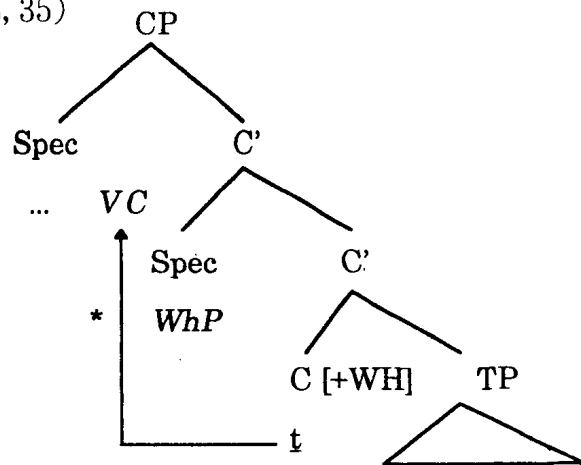
A wh-phrase (WhP) is attracted to the Spec of CP headed by C bearing [+WH]. The verbal complex (VC) must raise to a head which is higher than the head C in order to obtain the relevant word order. No such head exists, however. Thus, the forced raising of VC to a position higher than C is unmotivated, i.e., the *Economy Principle* violation.

The examples in (48a-c) are accounted for by the *Extension Condition,* as in (53), which is descriptively stated (Uriagereka 1998; 260)[36].

(53) *The Extension Condition*

The operation that applies at the first cycle must correspond to "inner" structures, while "outer" structures come next.

That is, given α and β, α commanding β, take first the lower β for an operation to apply, and then the higher α : the *bottom-up* solution. For (48a-b), the only assumption we need is that an adjunct-wh phrase *naze,* being adverbial in nature, merges with VP at the initial stage of the derivation. The rest follows from the *Extension Condition* and the *Overt-Wh-Category Movement Hypothesis.* The derivations of (48a-b) are (54a-b), respectively.

(54) a. (=48a)

```
                    CP
      Hanako-nom          C'
            what-acc          C'
                     why          C'
                            C       TP
                        eat-pst-Q  t    T'
                                       T    vP
                              8        t   t'    v'
                          7                    v      VP
                          6                         t    t    VP
                                                      4    t    V'
                                                                V    t
                                             3
                                                        1      t
                                                    2
```
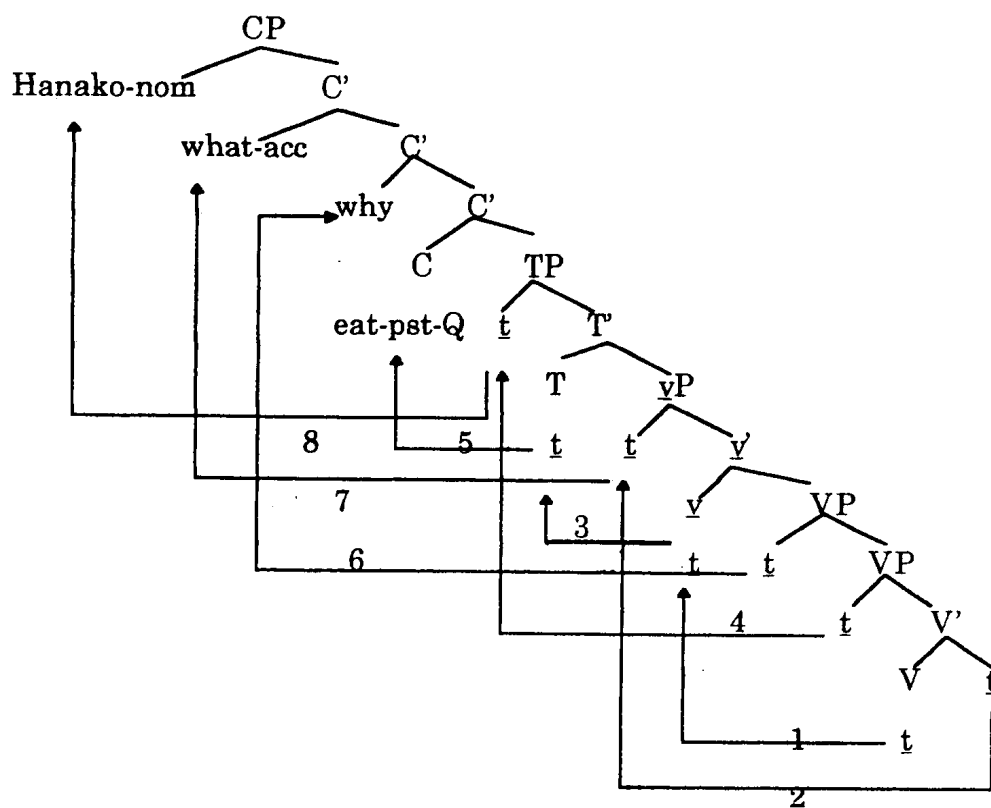
The number on an arrow indicates the order of steps of the derivation. One might argue that the derivation in (54a) does not obey the *Extension Condition* at all: a head H must raise first to extend the checking domain when some element lower than H must then undergo feature-checking at the Spec of H. This is required by the *Minimality Condition,* one of the *Economy principles.* Thus, a raising of a head H and the feature-checking of the lower DP at the Spec of H are immune from the *Extension Condition.* The *Extension Condition* sees cycles among DP's that are attracted by the higher head. Suppose we have completed Step 5. The next step is 6, in which the adjunct-wh phrase is attracted to C bearing [+WH]. Step 7 then follows, in which the direct object wh-phrase is attracted to C bearing [+WH]. Finally, the subject raises at Step 8. These applications

observe the *Extension Condition.* Consider the derivation of (48b).

(54) b．(=48b)
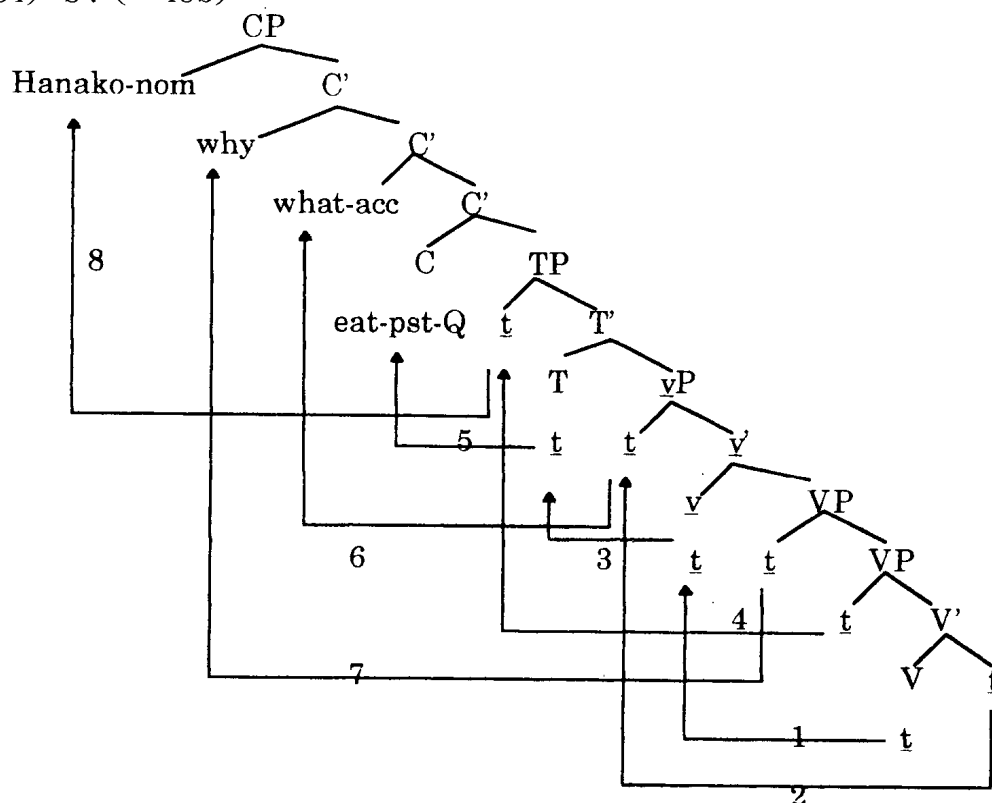
(54a) and (54b) share common steps up to Step 5. In (54b), the direct object wh-phrase raises at Step 6, which is followed by Step 7, in which the adjunct-wh raises to the Spec of C. Steps 6 and 7 violate the *Extension Condition,* thereby inducing the unacceptability of (48b). Notice that MLC is helpless: all source-targets are equidistant from C, given VC-movement to C.

Why does (48c) improve with a stressed adjunct-wh phrase? Our proposal is the following. When TP is formed in the derivation, the adjunct wh-phrase adjoins to TP for [+focus]-feature checking, thereby rendering the whole targeting cyclic[37]. The derivation of the example in (48c) is as follows.

(55) (=48c)
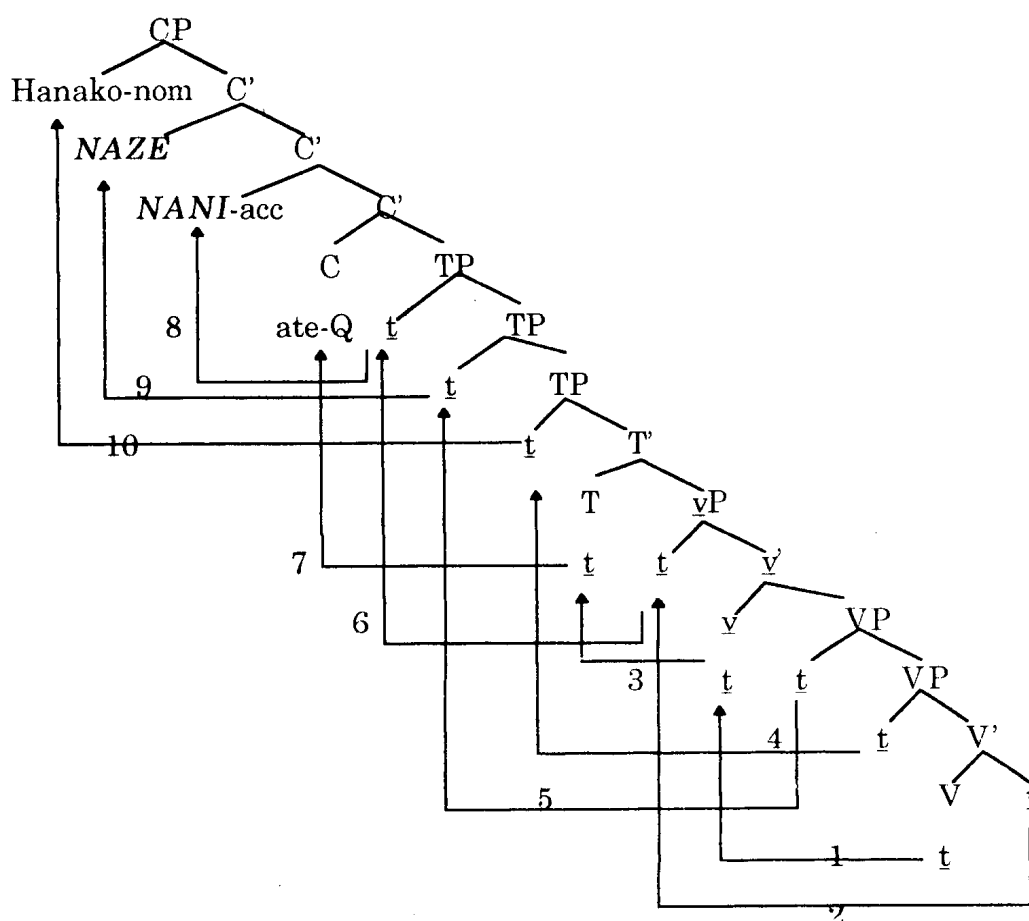
The steps that (54a-b) and (55) share continue to Step 4. At Step 5 and Step 6 in (55), the adjunct-wh phrase and the argument-wh phrase bearing [+FOCUS] adjoin to TP to check the relevant features. The verbal complex (VC) raises to C at Step 7. At Step 8, the direct object wh is attracted to C. The adjunct-wh raises to C at Step 9. The subject raises to C at Step 10. The last three steps do not violate the *Extension Condition,* since adjoined positions are equidistant to each other. Thus, the examples in (48a-c) are accounted for in a simple way[38].

6 ． A Simple FL-Solution: Argument vs. Adjunct-Wh Asymmetry

An argument-wh phrase can take the matrix scope, whereas an adjunct-wh phrase cannot.

(56) a . Hanako-ga [nani-o katta hito]-ni attano?

　　　　Hanako-nom what-acc bought person-dat met-Q

　　　　'(Lit.) Hanako met [the person who bought what]?'

　　b . * Hanako-ga [sono hon-o naze katta hito]-ni attano?

　　　　Hanako-nom　the book-acc　why　bought person-dat met-Q

　　　　'(Lit.) Hanako met [the person who bought the book why]?'

　　　　　　　　　　　　　　　　　　　　　　　　(cf. Huang1982)

The example in (56a) means "what is x, x a thing, such that Hanako met the person who bought x." However, the example in (56b) cannot mean "what is the reason y such that Hanako met the person who bought the book for y." The contrast disappears when the complex DPs in (56) are located post-verbally.

(57) a . * Hanako-ga attano [nani-o katta hito]-ni?

　　　　Hanako-nom met-Q what-acc bought person-dat

　　　　'(Lit.) Hanako met [the person who bought what]?'

　　b . * Hanako-ga attano [naze sono hon-o katta hito]-ni?

　　　　Hanako-nom met-Q why the book-acc bought person-dat

　　　　'(Lit.) Hanako met [the person that bought the book why]?'

These examples pose a problem for the ECP-based account. Given the *Stranding Approach,* the complex DP is stranded in the Spec of $\underline{v}$P after checking its dative-Case feature. There is no reason why the argument-wh phrase *nani-o* 'what-acc' in (57a) cannot raise to the Spec of the matrix CP, given that an argument-wh trace is properly-governed. The ECP-based account wrongly predicts that the example in (57a) should be acceptable, as in (56a)[39].

We propose that the *Wh-Pied-Piping Approach* (Nishigauchi 1986), which is slightly modified to accommodate the LCA (Kayne 1994), explains the examples in (56) and (57). First of all, the examples in (57a-b) are excluded as (6b, 7d-f, 8b): the complex DP is in the Spec of the matrix CP, thereby inducing the Economy principle violation if we force the verbal complex (VC) to a non-existent head higher than C. The derivation of (57a-b) are schematically indicated as follows[40].

(58)

* Economy principle violation
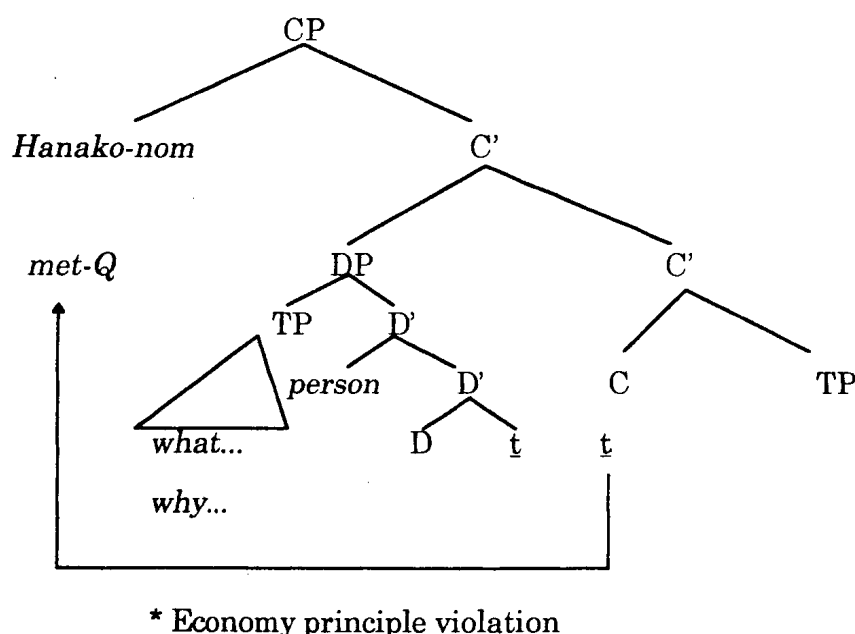
In (58), VC is forced to raise to a head higher than C. However, no such

head exists. Thus, the raising of VC is superfluous, inducing the Economy principle violation.

Now how do we account for the contrast in (56)? Let us repeat the examples.

(59) (=56) a. Hanako-ga [nani-o katta hito]-ni attano?

Hanako-nom what-acc bought person-dat met-Q

'(Lit.) Hanako met [the person who bought what]?'

b. * Hanako-ga [sono hon-o naze katta hito]-ni attano?

Hanako-nom the book-acc why bought person-dat met-Q

'(Lit.) Hanako met [the person who bought the book why]?'　　　　(cf. Huang1982)

We adopt the following conditions.

(60) a. [Wh]-feature of a wh phrase in an embedded clause is checked off by [wh]-feature of the matrix clause C only if the embedded clause is contained in the Spec of the matrix CP.

b. [V]-feature of an adjunct-wh phrase *naze* 'why' is checked off by C in the clause within which it is interpreted[41].

The intuitive idea behind the condition in (60a) is to weaken the Spec-head agreement. Thus, the following kind of Spec-head agreement is possible.
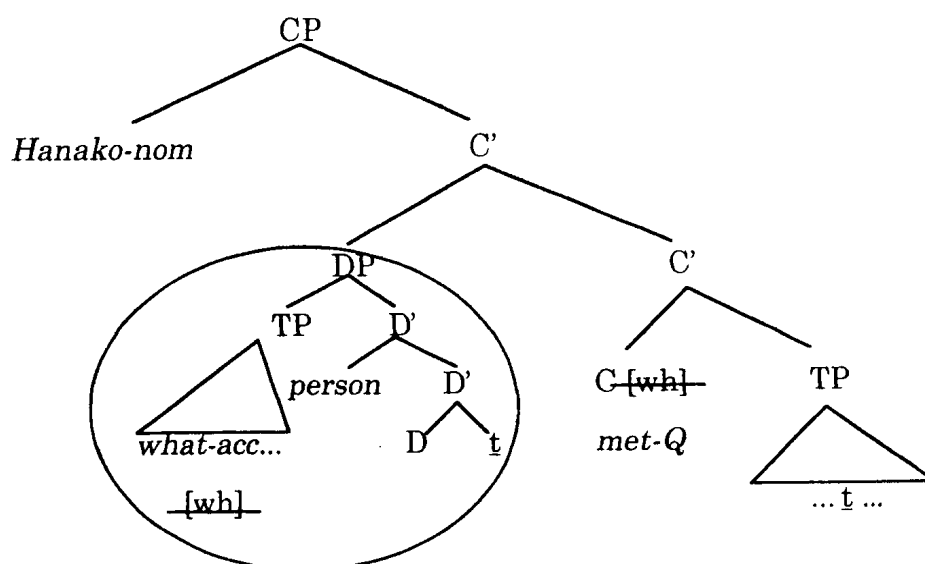
(61)

```
                    FP
                  /    \
                 α      F'
               / \     /  \
              /   \    F [f̶]  ...
             /_____\
           ... β [f̶] ...
```
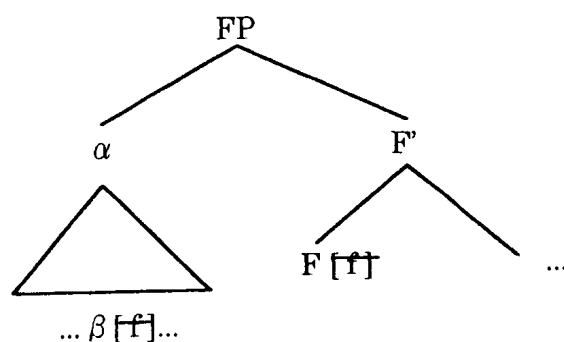
In (61), F is a functional category, α is the Spec of FP, [f] is an uninterpretable feature that must be erased. In (61), an uninterpretable feature [f] of β, which is contained in the Spec of FP, is checked off by being contained in α, which is the Spec of FP. The standard feature-checking under Spec-head agreement takes place between α and F. We dispense with feature-percolation analysis[42]. The derivation of the example in (59a) is shown in the following. We assume that a relative clause is a TP, not a CP[43].

(62)

```
                        CP
                      /    \
               Hanako-nom   C'
                           /    \
                          DP      C'
                         /  \    /   \
                       TP    D'  C̶-[w̶h̶]̶  TP
                      /\   person  D'   met-Q   /\
                     /__\        / \          /___\
                   what-acc...  D   t         ... t ...
                   [w̶h̶]̶
```
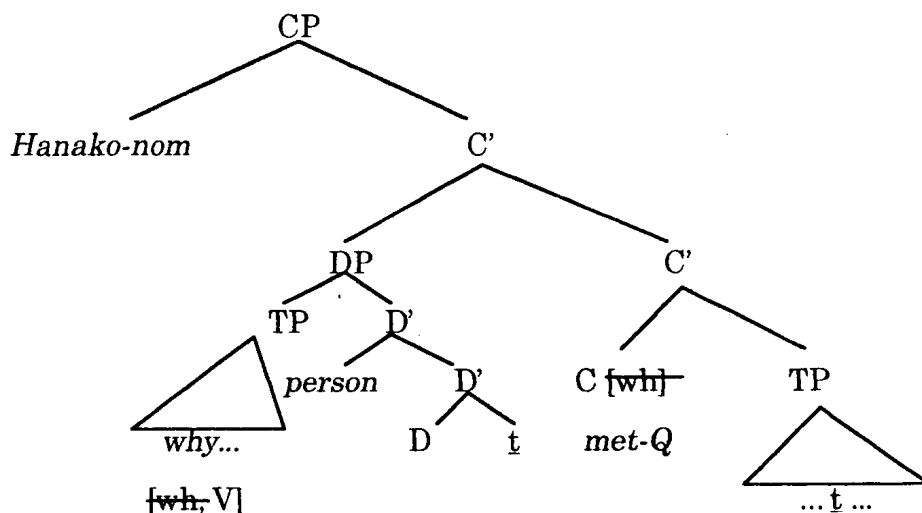
In (62), the [wh]-feature of the argument-wh phrase *nani-o* 'what-acc' is

checked off by the matrix C bearing [wh], a weaker version feature-checking under Spec-head agreement. Let us next consider the derivation of the example in (59b). Suppose that the derivation has reached the following stage.

(63)

By the condition in (60a), [wh]-feature of the adjunct-wh phrase *naze* 'why' is checked off by being in the Spec of the matrix clause CP headed by C bearing [wh]. By the condition in (60b), [V]-feature of the adjunct-wh phrase must be checked off by a local C within the embedded clause in which the adjunct-wh phrase is interpreted. However, no such C exists in the embedded clause. [V]-feature remains, causing a crash. The analysis predicts that an adjunct-wh phrase within an embedded CP should be permitted. The prediction is borne out.

(64) a . Hanako-ga [[cp Taro-ga naze sore-o katta] kara] okotteruno?

　　　　Hanako-nom Taro-nom why it-acc bought since is-angry

　　　　'(Lit.) Hanako is angry [since Taro bought it why]?'

b . Hanako-ga [[cp [Taro-ga naze sore-o katta] toiu] uwasa]-o kiitano?

Hanako-nom   Taro-nom   why it-acc bought that rumor-acc heard

'(Lit.) Hanako heard the rumor [that Taro bought it why]?'

In these example, the adjunct-wh phrase raises to the Spec of the embedded CP and checks [wh]-feature off. The derivation of (64a) is shown in the following[44].

(65)
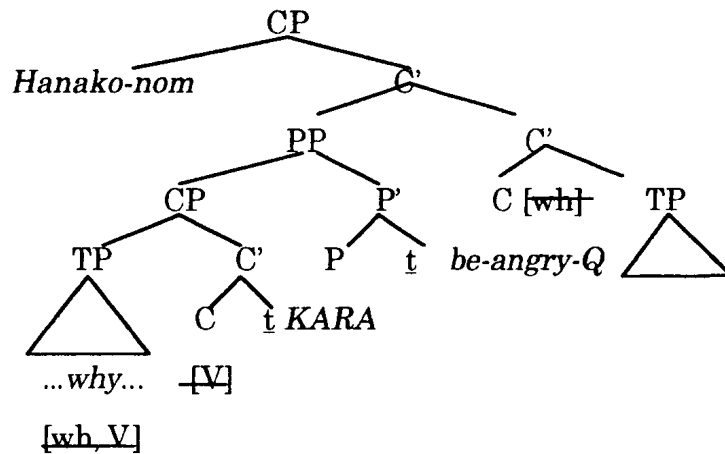
In (65), [wh]-feature is checked off by Spec-head agreement. [V]-feature is checked off by the local C. Consider the derivation of the example in (64b).

(66)

```
                        CP
                  ┌──────┴──────┐
           Hanako-nom           C'
                          ┌──────┴──────┐
                         DP             C'
                   ┌──────┴──────┐   ┌───┴────┐
                  CP            D'  C-[wh]    TP
               ┌──┴──┐    ┌────┴────┐        /\
              TP    C'  rumor      D'       /  \
              /\   ┌┴┐         ┌────┴┐     /____\
             /  \  C  t        D     t
            /____\
            ..why..   toiu
            [wh, V]    [V]
```
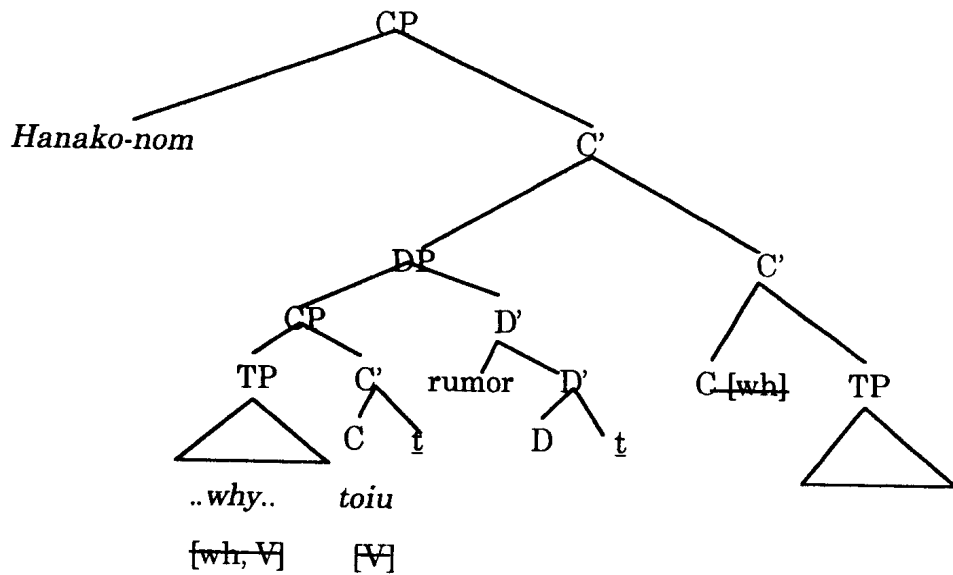
In (66), the complex DP formation proceeds as follows. D and CP merge. The nominal *rumor* merges with D'. The remaining CP adjoins to D'. [wh]-feature is checked off by Spec-head agreement. [V]-feature of the adjunct-wh phrase is checked off by the local C in the embedded clause within which the adjunct is interpreted. Thus, the argument-adjunct wh asymmetry is accounted for by the *Overt-Wh-Category Movement Hypothesis* and the *Wh-Pied-Piping Approach*. The ECP-based accounts are dispensed with.

We have observed that an adjunct-wh phrase *naze* 'why' is permitted within the embedded clause which has the status of CP. However, the example becomes unacceptable when the adjunct-wh phrase is overtly scrambled out of the embedded CP, which needs an explanation. Consider the following contrast. The grammatical judgments are Sohn's (1994).

(67) a . *naze John-wa [[cp Mary-ga t sono hito-o uttaeta to iu] uwasa]-o kiita-no

　　　　 why John-top　　Mary-nom the man-acc sued that rumor-acc heard-Q

　　　　 'Why₍ did John hear [the rumor [that Mary sued the man t₍]]'

b. ?(?) dare-o John-wa [[cp Mary-ga t̲ uttaeta to iu] uwasa]-o kiita-no

　　　who-acc John-top Mary-nom　sued that rumor-acc heard-Q

　　　'Who$_i$ did John hear [the rumor [that Mary sued t̲$_i$]]'
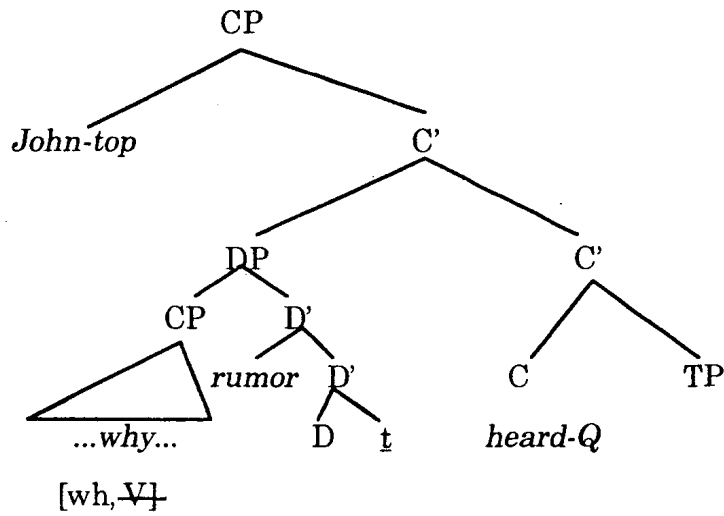
　　　　　　　　　　　　　　　　　　　　　　(Sohn 1994)

The standard account is that the example in (67a) violates both the ECP and Subjacency condition, whereas that the example in (67b) violates the Subjacency condition alone. Since we have dispensed with the ECP, an alternative is called for. We propose that the contrast is accounted for by the unsaid assumption (68) that is implicit in the condition (47) (*Universal WH-Licensing Condition*), which we have been assuming throughout.

(68) A wh-phrase without agreement-features raises to the Spec of CP, unless the movement violates the Economy conditions.

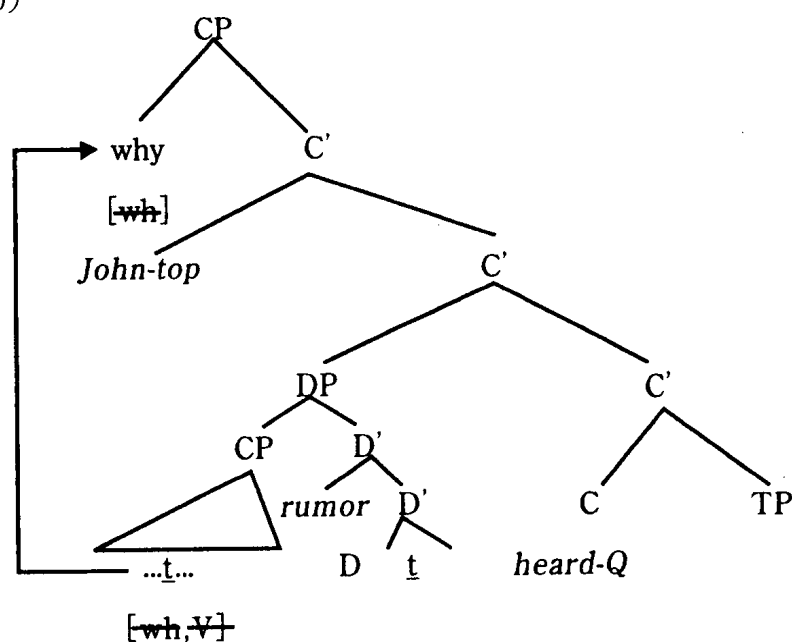One of the cases in which a wh-phrase lacks agreement-features is the case of an adjunct-wh phrase. (68) states that all adjunct-wh phrases must raise to the Spec of the relevant CP to be interpreted, whereas a non-agreeing wh phrase can be Q-bound if it is contained in the Spec of C bearing [wh]-feature, i.e., (61). Let us consider the derivation of (67a). Suppose that the derivation has reached the following stage.

(69)

```
                        CP
                   ┌─────┴──────┐
              John-top          C'
                          ┌──────┴────────┐
                         DP               C'
                     ┌────┴──┐         ┌───┴────┐
                   CP        D'        C        TP
                 ┌──┴┐    ┌──┴──┐
                 ...    rumor  D'
              ...why...      ┌──┴─┐   heard-Q
               [wh,V̶]        D    t
```

The complex DP is in the Spec of the matrix CP. Suppose that the uninterpretable [V]-feature of the adjunct-wh phrase is checked off within the embedded clause, in which the phrase is interpreted, but that the uninterpretable [wh]-feature is not erased at this point. The adjunct-wh phrase must raise to the Spec of CP to check [wh] off. The relevant structure at this point is as follows.
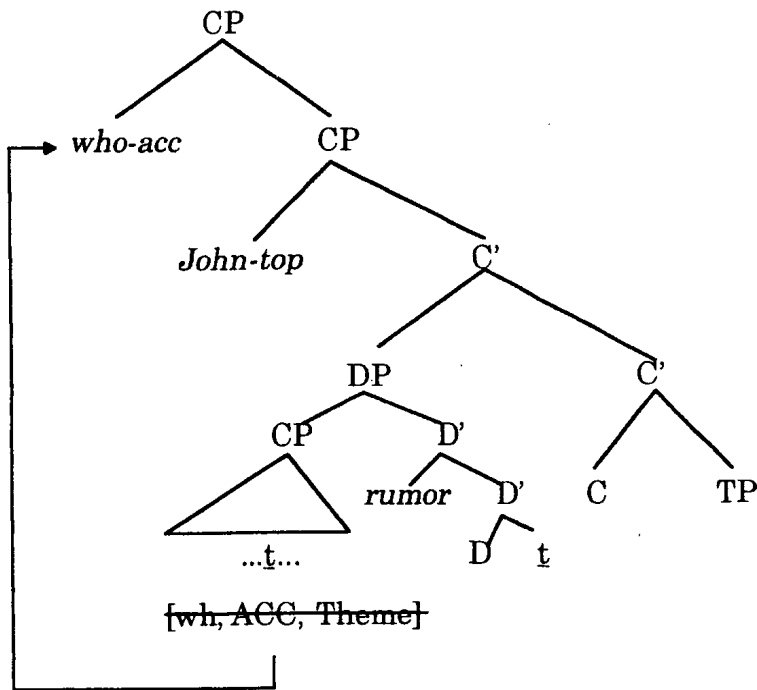
(70)

```
                              CP
                            /    \
              ┌──→  why       C'
              │      [wh]    /    \
              │     John-top      C'
              │              /        \
              │            DP          C'
              │          /    \       /   \
              │        CP      D'    C     TP
              │       /  \  rumor  D'
              │      / △  \      /  \
              └── ...t...     D   t   heard-Q
                  [wh,V]
```

In (70), since the adjunct-wh phrase is in the Spec of CP, it agrees with the head C in the matrix clause. That is, the adjunct-wh phrase is interpreted as the matrix adjunct, i.e., what is the reason y, such that John heard for the reason y [the rumor that Mary sued the person]. This is not the interpretation we want for (67a). Thus, all the features of the adjunct-wh phrase [wh, V] must be erased within the embedded clause. Alternatively, if we assume that all features of *naze* are erased within the lower CP, nothing forces further movement. If it is forced to raise to the Spec of the matrix CP, as in (67a), the movement is in violation of the Economy principle, which prohibits a superfluous (unmotivated) movement.

What about the example in (67b)? Why is it relatively acceptable? We propose that it is accounted for by the condition in (68). Suppose all features of the argument-wh phrase *dare-o* 'who-acc' are erased within the embedded clause. The derivation of the example in (67b) is as follows.

(71)

```
                    CP
                  /    \
        who-acc          CP
                       /    \
              John-top        C'
                            /    \
                       DP          C'
                     /    \       /   \
                  CP        D'    C     TP
                 /  \     /    \
               /____\  rumor   D'
               ...t...        /  \
                             D    t
            [wh, ACC, Theme]
```

Thus, the raising in (67b) is a scrambling operation, not a feature-checking[45]. A scrambling operation is possible for an argument-wh phrase, since there is no requirement that the argument must be in the Spec of CP. We assume that the landing site of scrambling is an adjoined position, not a Spec. It follows that we still need to make the distinction between an adjoined position and a Spec, contra Kayne (1994) and in accordance with Chomsky (1994); only the latter involves feature-checking. The distinction is schematically indicated as follows, where F = a functional category.

(72) a. Feature-Checking   b. Scrambling
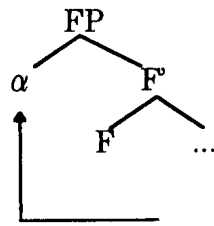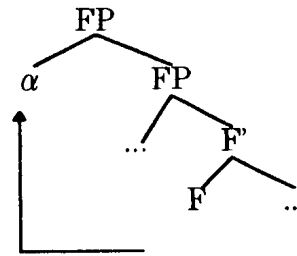
α in (72a) must be interpreted at this place, whereas α in (72b) is not interpreted at this position. In (67a), the adjunct-wh phrase is in the Spec of CP and must be interpreted at the matrix-clause level, which leads to a contradiction to the expected reading, whereas in (67b), the argument-wh phrase is adjoined to CP and it is not interpreted at the matrix-clause level. Only an argument-wh phrase can undergo scrambling without checking of its [wh] at the matrix-clause level.

7. Concluding Remarks

Let us itemize the points. First, we have discussed the theory of word order called LCA, and have reached the following conclusion.

(73) The LCA is respected in the human language (= section 3).

The *Linear Correspondence Axiom* (LCA) (Kayne 1994) states that the higher a terminal element occupies in the Command configuration, the earlier the term occupies in the precedence relation. We have provided evidence from Hindi and Japanese that LCA is working in C$_{HL}$, the computational procedure for human language[46]. We then reached the following conclusion.

(74) Any wh-phrase without agreement-features raises to the Spec of CP unless some Economy principle forces it not to (= section 4).

We have observed that FL solved the Legibility Problem with respect to wh-movement in a simple, elegant, and natural way: all wh-phrases overtly raise to the Spec of CP. However, the agreement system overrides the FL-solution: if a wh-phrase bears [+AGR]-feature, it is frozen in the Spec of the feature-checking position. Since such wh-phrase cannot raise to the Spec of CP, it must be Q-bound by C bearing [+Q], i.e., the Economy condition of the Last Resort version. We then observed two wh-related phenomena in the light of the *Overt-Wh-Category Movement Hypothesis*. We first saw the multiple-wh effect in Japanese, and reached the following conclusion.

(75) The multiple-wh effect is a well-behaved cyclic application (section 5).

Given our hypothesis, the multiple-wh effect is accounted for by the Extension Condition alone. No stipulated conditions such as Anti-Superiority Conditions or LF-Saving Effect Condition are necessary. Furthermore, if a cyclic application is the only option the design of C_{HL} has to realize its negentropical nature, the multiple-wh effect is explained by the Economy conditions alone, given the *Overt-Wh-Category Movement Hypothesis*[47]. We then observed the argument-adjunct-wh asymmetry within the islands, and reached the following conclusion.

(76) The facts of argument-adjunct contrast of wh-movement support the *Overt Wh-Category Movement Hypothesis* (section 6).

Given these observations, we have reached the following conclusion.

(77) The FL solution to the Legibility Problem is optimal.

Below are remaining problems, which we leave for the future research.

(78) a . How solid is the argument that $[\pm \text{AGR}]$ is subject to the parametric variation? Does Japanese lack $[+\text{AGR}]$ entirely?

   b . How general is the observation that agreement-feature checking terminates further movement? Is the formulation of *Greed* valid?

Notes

1) The research reported here was supported in part by grants-in-aid for Encouragement of Young Scientists from the Japanese Ministry of Education, Science, Sports and Culture (1998-1999; Project #: 10710260; Title: *"A preliminary study of syntactic-feature representation in human language and its implications for brain sciences"*)

2) For similar proposal, see Chomsky (1995, 1998) and Epstein at al. (1998). We follow the basic line of approach referred to as the Minimalist Program (MP). MP has not yet attained the status of "theory." For basic assumptions of MP, see Chomsky (1995, 1998) and Uriagereka (1998, 1999).

3) See Fodor (1983) for the concept "a language of thought." For AP ∕ CI distinction of FP, see Chomsky (1995; 168). The theory of FL at its initial state S_0 is called the Universal Grammar (UG). The theory of "attained state" is called Particular Grammars. The study of language organ (FL) is similar to the study of other organ systems as the visual system, immune system, or circulatory, system in the sense that we cannot extract each
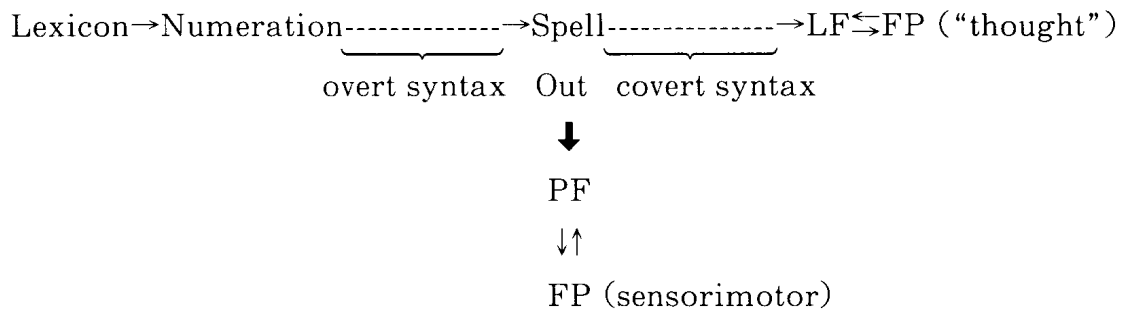
system from the whole. For the relevant assumptions and speculations adopted in this section, see Chomsky (1995, 1997, 1998).

4) That is, it exerts its influence on the phonetic side.

5) More precisely, all strong uninterpretable features must be erased by Spell-Out. Weak uninterpretable features can be erased after Spell-Out. See Chomsky (1995) for strength of features. However, all uninterpretable features must be erased by the time when the derivation reaches LF, since LF is the interface-level, where all features must be interpreted, i.e., be legible to FP.

6) The overall picture of FL-FP relation is visualized as follows.

(i)

Lexicon→Numeration-------------→Spell---------------→LF⇆FP ("thought")
　　　　　　　　　⎵‾‾‾‾‾‾‾‾‾‾‾‾⎵　Out　⎵‾‾‾‾‾‾‾‾‾‾‾⎵
　　　　　　　　　　overt syntax　　　　covert syntax
　　　　　　　　　　　　　　　　↓
　　　　　　　　　　　　　　　　PF
　　　　　　　　　　　　　　　↓↑
　　　　　　　　　　　　FP (sensorimotor)

Lexicon is the component of FL in which lexical items (both functional and substantive) are stored. Numeration is the component of FL where necessary lexical items are selected. Once a derivation leaves the Numeration, no further access to it is permitted (Inclusiveness Condition; Chomsky (1995)). The Inclusive Condition is "some kind of Conservation Law that tells you something to the effect that whatever chains you have at LF (and only those) are the result of manipulating whatever lexical features you had in your numeration" (Uriagereka 1999; 59). See Uriagereka (1995) and Epstein at al. (1998) for an alternative hypothesis (a strong derivational approach), in which FP has direct access to the computation itself, i.e., Spell-Out applies at any stage of the derivation (multiple Spell-Out).

7) In fact, it is an operator-variable relation: Q (operator) is licensed when there is a wh-phrase (variable) that is bound by Q, avoiding vacuous quantification. See Heim (1982), Pesetsky (1987), Nishigauchi (1990), Tsai (1994), and Saito (1998).

8) A dog may have FL, but her FL is not legible to her FP, which is a logical

possibility (Chomsky 1997). In this connection, Chomsky (1998; 6-7) provides an evolutionary fable, which shows the relation among FL, FP, and the Legibility Problem.

> "...Imagine some primate with the human mental architecture and sensorimotor apparatus in place, but no language organ. It had our modes of perceptual organization, our propositional attitudes (beliefs, desires, hopes, fears, ...) insofar as these are not mediated by language, perhaps a "language of thought" in Jerry Fodor's sense, but no way to express its thoughts by means of linguistic expressions, so that they remain largely inaccessible to it, and to others. Suppose some event reorganizes the brain in such as way as, in effect, to insert FL. To be usable, the new organ has to meet certain "legibility conditions." Other systems of the mind/brain have to be able to access expressions generated by states of FL ((I-) languages), to "read" them and use them as "instructions" for thought and action. We can try to formulate clearly -- and if possible answer -- the question of how good a solution FL is to the legibility conditions, and these alone. That is essentially the topic of the minimalist program."

9) For the definition of LCA, see section 3.1.

10) It should be emphasized that it is not a finished product (theory) at all. We are just beginning to formulate questions as to what the biological bases of human language are.

11) Structural-Case features (formal features) are typical uninterpretable features. They are uninterpretable because they do not contribute to semantic or phonetic interpretation. Consider the following examples.

(i) John-nom hit Mary-acc

(ii) Mary-nom was hit by John

In (i), the object DP *Mary* has THEME (something affected) interpretation, and has acc-Case. In (ii), the subject *Mary* still has the THEME interpretation, but it has nom-Case. If structural Case contributed to meaning, Case of *Mary* in (i) and (ii) should be identical. The fact that the two Cases are distinct indicates that Case does not contribute to the meaning. Furthermore, formal features are uninterpretable at PF. Thus, they are not

interpreted at any interface. The other characteristics of Cʜʟ are discrete infinity (the property of natural numbers: there are three-word sentences and four-word sentences, but not three-and-a-half-word sentences or 4.7-word sentences) and plasticity (FL grows (and dies) like other organs).

12) [wh] is an uninterpretable, whereas [Q] is interpretable (Chomsky 1998).

13) That is, human language is not the result of adaptation. To quote from Uriagereka (1998; 48-49), which adopts the notion of *exaptation* (Stephen Jay Gould 1991);

"... the language faculty might have arisen in the human species as a result of properties of the brain that evolved for entirely different reasons, having nothing to do with the language faculty per se. ... language might be an evolutionary *exaptation,* which is a feature that may be useful for a given organism but didnt originate as an adaptation to fulfill its current function ..."

That is, the adaptationist view is that FL emerged because of some functional need, whereas the exaptationist views is that FL is the by-product of some other form whose origin may or may not itself be as an adaptation (e.g., the brain grew larger for some reason, and that liberated enough 'mind space' to allow for linguistic structures) (Uriagereka 1999). Given the initial state of FL, which is genetically determined and constrained by certain biochemical properties of human brain, human language is more like a genetic expression of a mutation in which FL somehow became able to access and interact with FP. See Chomsky (1995, 1998) and particularly Uriagereka (1998). Uriagereka (1998) speculates that human language is just another example of the realization of workings of the Fibonacci series, which we find everywhere in the natural world, e.g., sunflower seeds, tree branches, and shell formation.

14) S＝subject, O＝object, V＝verb, J＝Japanese, H＝Hindi, nom＝nominative Case, acc＝accusative Case, dat＝dative Case, erg＝ergative Case, abs ＝absolute Case, m＝masculine, fem＝feminine, perf＝perfective.

15) I am indebted to Kevin R. Gregg for methodological improvement and detailed information. SOV order in (5a) can be found in poems (*John an*

apple ate). (5c) is acceptable if *Bill* is topicalized (*Bill, John killed*). The marginal status of (5e) may be connected to the fact that the V-initial order can be found for yes／no questions in old English.

16) The judgment ??? of the relevant example is from Mahajan (1997). The judgment should be understood in a relative sense, which means that the status ??? indicates that something is wrong within the derivation of that example.

17) For the definition of the MLC, see Chomsky (1995) and Kitahara (1997). The MLC is subsumed under the definition of Attract (Uriagereka 1998; 537-538).

(i) *Definition of Attract* ($\supset$ the MLC)

K *attracts* F if F is the closest feature that can enter into a checking relation with a sublabel of K.

(ii) *Definition of Distance*

If β commands α, when targeting K for raising, with τ the actual target of movement, β is *closer* to K than α is, unless

(a) β and τ are in the same minimal domain, or

(b) α and β are in the same minimal domain.

(iii) *Definition of minimal domain*

Where α is a feature matrix or a head #X#, and CH is a given chain (α, t) or (the trivial chain) α:

a. Max (α) is the smallest maximal projection dominating α.

b. The domain D (CH) of CH is the set of categories／features dominated by Max (α) that are distinct from and do not contain α or t.

c. The *minimal domain* Min (D (CH)) of CH is the smallest subset K of D (CH) such that for any x belonging to D (CH), some γ belonging to K dominates x.

18) *Command* was previously referred to as *C-command*. The original proposal of the LCA made in Kayne (1994) is the following. Let us call d(X) the set of terminals that X dominates. Given a phrase marker P, let T the set of terminals in P, and S be the maximal set of ordered pairs $<$X, Y$>$ where X, Y are nonterminals in P and X asymmetrically c-commands Y. Then,

(i) *Linear Correspondence Axiom*

　　　d(S) is a linear ordering of *T*.　　　　(Kayne 1994; 5-6)

The definition of *c-command* adopted in Kayne (1994) is the following.

(ii) X *c-commands* Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y.

(iii) α *dominates* β iff every segment of α dominates β.

(iv) α *excludes* β iff no segment of α dominates β.

19) For our purpose, IP＝AGRsP＝TP, and AGRoP＝AGRdoP＝$\underline{v}$P. Nominative Case is checked at the Spec of TP, and accusative Case is checked at the Spec of $\underline{v}$P.

20) The terminology A and A' are used to indicate that there are two types of movement.

21) The *Stranding Approach* can explain, without complication, the fact that the post-verbal object QP agrees with the verb.

(i) Mohan-ne Ravi-ko $\underline{t}$ bhejii [har ek kitaab]

　　Mohan-erg Ravi-dat sent-fem every book-fem

　　'Mohan sent to Ravi every book'　　　(Mahajan 1997)

The verb adjoins to AGRo on its way to a higher head. When the verb is under AGRo, the QP *har ek kitaab* 'every book' is in the Spec of AGRoP, where its Case is checked off, thereby showing the object agreement.

22) In Japanese as well, various tests such as the *Additional-WH test, Binding tests, Bound Variable Binding test, SCO test, Idiom Chunk test, Locality test, NQ test, Parasitic Gap test, Reconstruction tests, Scope test,* and *Stress test,* but not the *WCO test,* show that a leftward-long-scrambled phrase moving out of a finite clause occupies a position bearing A'-properties.

23) There is a peculiar asymmetry between Hindi and Japanese with respect to WCO amelioration when long-distance scrambling is involved: In Hindi, a long-distance-scrambled non-wh-QP contributes to WCO remedy, but a

long-distance-scrambled wh-phrase does not, whereas in Japanese, both non-wh-QP and wh-phrase contribute to WCO remedy.

(i) a . har ek kitaab₁ Gita-ne socaa [ki Mohan-ne us₁-ke maalik-ko bhejii]

every book-acc Gita-erg thought that Mohan-erg his owner-dat sent

'Every book, Gita thought that Mohan sent to its owner'

b . *kis₁-ko us₁-ki bahin-ne socaa [ki Ram-ne t̲₁ dekhaa thaa]?

who-acc his sister-erg thought that Ram-erg saw AUX

'Who₁ did his₁ sister think that Ram saw?'   (Mahajan 1990; 39)

(ii) a . daremo1-o Hanako-ga [Taro-ga soitu1-no hahaoya-ni hikiwatasita to] itta

everyone-acc Hanako-nom Taro-nom his mother-dat introduced that said

'Everyone, Hanako said that Taro introduced t̲ to his mother'

b . ? dare₁-o soitu₁-no hahaoya-ga [Hanako-ga t̲₁ aisiteiru to] itta no?

who-acc his mother-nom Hanako-nom love that said Q

'(Lit.) Who did his mother say that Hanako loved?' (Yoshimura 1992)

24) The theory predicts the following: if the indirect object (IO) is QP, and the direct object (DO) contains a variable, the sentence would be acceptable. The prediction is borne out.

(i) Mohan-ne har ek aadmii₁-ko loTaaii [us₁-kii kitaab]

Mohan-erg every person-dat returned his book

'(Lit.) Mohan returned to every person, his book'   (Mahajan 1997)

(ii) Taro-wa daremo₁-ni kaesita [soitu₁-no hon-o]

Taro-top everyone-dat returned his book-acc

'(Lit.) Taro returned to everyone, his book'

25) AGR₁o checks the Case of indirect object (Mahajan 1997).

26) We do not assume [+WH]-feature movement. See Chomsky (1998).

27) Our analysis supports the hypothesis that [±AGR] is subject to the parametric difference, contrary to the AGR-based theory adopted in

Watanabe (1993) and Koizumi (1995).

28) This predicts that the following example in Hindi should be acceptable with Q-binding.

(i) kyaa ciiz Ram-ne khaaii?

what thing-fem Sita-erg ate-fem

'What did Ram eat?'

In (i), *kyaa ciiz* is in the Spec of AGRoP, *Ram-ne* is in the lower Spec of AGRoP, and the verb *khaaii* is adjoined to AGRo head.

29) The relevance of agreement-features, but not of Case-features, is supported by the possibility that the formatives *who* and *what* do not bear Case-features, given the following examples.

(i) who do you wonder whether Mary loves him?

(ii) who do you wonder whether Mary loves John?

In (i) and (ii), *who* does not check Case-feature. No principle is violated. Thus, the examples converge, and they are legible, but not intelligible (Chomsky 1998, Uriagereka 1999). Furthermore, the total integration of agreement-features and Case-features cannot be possible, contrary to what has been proposed in Chomsky (1998). According to Chomsky, feature-driven movements are divided into two types: direct feature-driven movements (DFM) and indirect feature-driven movements (IFM). An example of DFM is raising to subject (EPP). IFM is further divided into the following two categories: A-movement, in which the attracting head H bears ϕ-features of the Case-agreement system, and A'-movement, in which H bears P-features of the peripheral system (force, topic, focus, etc.).

30) [+AGR] is a member of the set of ϕ-features.

31) An expletive *there,* contrary to an expletive *it,* bears only the categorial feature D (EPP-feature). In the expletive-*there* construction in (42b), if the associate did not raise and adjoin to *there* covertly, Case and [+AGR] (ϕ) features of the matrix [Infl-*seem*] would not be checked off. Thus, in (42b), the associate has in fact undergone movement in the covert syntax and is adjoined to *there.*

32) The ungrammaticality of this example is the crucial empirical evidence for the LF crash theory, i.e., one of "three minimalist approaches to overt movement" (Chomsky 1995, Lasnik 1999; 198). Their example is the following.

(i) *John read what?

What induces the ungrammaticality is as follows. In order to be interpreted at LF, the derivation of (i) covertly introduces C with a strong feature [+wh] (a lexical material can be added as long as the addition takes at the root). However, [+wh] can only be checked off by overt wh-movement, which is too late for (i). Thus, the strong feature [+wh] of C remains at LF, inducing a crash at LF. This account, however, wrongly predicts that the following example is grammatical.

(ii) *Who did read it?

Assuming *did* occupies C, a strong feature [+wh] of C is checked off by overt wh-movement to the Spec of CP. The sentence should converge. Also unclear is why Q-binding is unavailable in (i). Thus, the LF crash theory is not tenable.

33) This indicates that Chomsky's (1998) "no-look-ahead" condition is too strict.

34) For evidence of overt movement of V to T in Japanese, see Koizumi (1995).

35) For mechanics of the ECP, see Chomsky (1981) and Lasnik and Saito (1992). The ECP as such does not exist in the Minimalist Program.

36) Originally proposed as *Strict Cyclicity* in Chomsky (1973). Watanabe (1995) discusses the conceptual basis of Cyclicity. He argues that cyclic application is forced by the condition called Avoid Redefinition. Every node α, which dominates a newly created node in the course of derivation, is redefined. In the cyclic derivation, the number of α decreases as the derivation proceeds. In the countercyclic derivation, the number of α increases as the derivation proceeds. The greater the number of α, the more redefinitions occur, which means the more "disturbances" there are to the legibility problem. Thus, the cyclic application is required to minimize (= economize) the disturbances caused by Move or Merge. Although Watanabe

(1995) discusses only the target cyclicity, the discussion holds both for the target and the source of application. The analysis presented in this section provides empirical evidence for Avoid Redefinition Condition.

37) See Nunes (1995) for the proposal that the extention condition applies to Adjunction as well as Merge and Move.

38) Sohn (1994) reports the following contrast.

(i) a. ?(?) dare-o1 naze2 John-wa [[Mary-ga t2 t1 uttaeta to iu] uwasa]-o kiita-no

     who-acc why John-top Mary-nom   sued   that rumor-acc heard-Q

     'Q John heard [the rumor [that Mary sued who why]]'

  b. ?? naze2 dare-o1 John-wa [[Mary-ga t2 t1 uttaeta to iu] uwasa]-o kiita-no

     why who-acc John-top Mary-nom   sued   that rumor-acc heard-Q

     'Q John heard [the rumor [that Mary sued who why]]'

Sohn (1994) judges (ib) to be slightly worse. However, we argue that (ib) is unacceptable (*) without stress. The same contrast appears here as the one we find in the matrix clause. The examples supports our analysis.

39) The ECP (Empty Category Principle) has no place in the MP (Minimalist Program) (Chomsky 1995, 1998, Uriagereka 1998). The standard definition of the ECP states that a nonpronominal trace must be properly governed (Lasnik and Saito 1992). The reason the ECP has been discarded is that the notion of government is too ugly. That is, the notion of government contains the following heterogeneous configurations.

(i) A head H governs the complement CPL.

(ii) H governs the specifier SPEC.

(iii) H governs the SPEC of the CPL (= Exceptional Case Marking)

(iv) H governs the subject of a small clause (SC)

It is hard to see the common property of CPL, SPEC, SPEC of CPL, and the subject of SC. There is no principled reason as to why it is 'exceptional' when H governs the SPEC of the CPL. The notion of government collects heterogeneous sets of different kinds, which suggests that the notion is unmotivated and ugly. Any unmotivated and ugly notion is discarded in the MP.

40) A complex DP (a relative clause +NP) formation proceeds as follows. First, D and TP merge. A relative head raises to the Spec of DP. The rest of the remaining TP raises to the Spec of DP. The hypothesis that a relative clause is a TP, not a CP, will be crucial in later discussion.

41) See Nishigauchi (1986) for a similar proposal. Reinhart (1993) and Tsai (1994) argue that the familiar argument-adjunct distinction is in fact an argument-adverb distinction, based on the fact that adverbs lack a position for a variable so that the in-situ interpretation is barred (Chomsky 1995; 386, fn. 65). The idea that an adjunct-wh phrase *naze* 'why' in Japanese is an adverb in nature is not new. Mikami (1953) has noticed that there are four types of wh-phrases as shown in the following.

(i) a. wh-nouns: *nani* 'what,' *dare* 'who,' *doko* 'what place,' *what.*

 b. wh-pronouns: *dore* 'which one,' *dare* 'who,' *which, who.*

 c. wh-adnominals: *dono* 'which,' *donna* 'what kind of,' *what.*

 d. wh-adverbials: *doo* 'how,' *itu* 'when,' *naze* 'why,' *when, where, why, how.*

42) See Nishigauchi (1986).

43) See Murasugi (1991). Kayne (1994) assumes that it is a CP.

44) As for evidence for CP-status of *KARA*-clause, see Arikawa (1998).

45) A scrambling operation is a problem in terms of Economy considerations. Given the Economy principle, every movement must have a reason to move. If we assume that the argument-wh phrase in (68b) moves to check some kind of [FOCUS]-feature, we then have to assume that an adjunct-wh phrase cannot bear [FOCUS]. We also have to say that [FOCUS] can be checked either by T or C. The issue is not clear.

46) For evidence for the LCA from other languages such as English and French, see Kayne (1994).

47) Negentropy is a concept proposed by Erwin Schrödinger to explain a system that does not obey the law of entropy. For example, a system of life must decrease the level of disturbances (mess) constantly in order for the system to keep working. If Watanabe (1995) is right in that CHL obeys the Avoid Redefinition condition, the cyclic nature observed in the computational procedure for human language is one of the negentropical workings

to decrease the disturbances of the derivations. Thus, the design of human language exhibits properties of inorganic systems such as Economy, as Noam Chomsky emphasizes, but it also shows properties of organic systems such as negentropy.

References

Arikawa, K. 1998. Modal Feature Erasure and the Principle Last Resort － Evidence for Another Formal Feature of C^0 －. Human Sciences Review No. 15. St. Andrew's University.

Chomsky, N. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, ed. Anderson, S.R. and Kiparsky, P., Holt, Rinehart and Winston, New York, 232-86.

Chomsky, N. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. 1994. Bare phrase structure. MIT Occasional Papers in Linguistics 5. MITWPL, Department of Linguistics and Philosophy, MIT. (Also published in *Government and Binding Theory and the Minimalist Program,* edited by G. Webelhuth. Oxford: Blackwell (1995).)

Chomsky, N. 1995. *The Minimalist Program.* MIT Press, Cambridge (*MP*).

Chomsky, N. 1997. Current Issues in Linguistics. A lecture given at Yale University.

Chomsky, N. 1998. Minimalist Inquiries: the Framework. ms. MIT.

Epstein, S., E. Groat, R. Kawashima, and H, Kitahara (1998). *A Derivational Approach to Syntactic Relations.* Oxford University press.

Fodor, J. 1983. *The Modularity of Mind.* MIT Press, Cambridge.

Gould, S. J. 1991. Exaptation: A crucial tool for evolutionary psychology. *Journal of Social Issues* 47, 43-65.

Kayne, R. 1994. *The Antisymmetry of Syntax.* MIT Press, Cambridge.

Heim, I. 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral dissertation, University of Massachusetts at Amherst.

Huang, C.-T. J. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge.

Kitahara, H. 1997. *Elementary operations and optimal derivations.*

Linguistic Inquiry Monograph 31. MIT Press.

Koizumi, M. 1995. Phrase structure in Minimalist syntax. Doctoral dissertation, MIT.

Lasnik, H. and Saito, M. 1992. *Move alpha.* Cambridge, Mass.: MIT Press.

Lasnik, H. 1999. On Feature Strength: Three Minimalist Approaches to Overt Movement. *Linguistic Inquiry* 30, 197-217.

Mahajan, A. 1990. The A∕A-Bar Distinction and Movement Theory. Doctoral dissertation, MIT, Cambridge.

Mahajan, A. 1997a. Against a Rightward Analysis of Extraposition and Rightward Scrambling in Hindi. In Tonoike, S. 1997. *Scrambling.* Linguistics Workshop Series 5, 93-124. Kuroshio, Tokyo.

Mahajan, A. 1997b. Rightward scrambling. In *Rightward Movement,* ed. Beerman, D., LeBlanc, D., Riemsdijk, H. v., John Benjamins, Amsterdam ∕ Philadelphia, 185-213.

Mikami, A. 1953. *Gendaigohoo zyosetu − Syntax no kokoromi −* "An introduction to modern language usage − An attempt for Syntax −." Republished by Kuroshio publication, Tokyo.

Murasugi, K. 1991. Noun phrases in Japanese and English: A study in syntax, learnability and acquisition. Doctoral dissertation. University of Connecticut.

Nishigauchi, T. 1986. Quantification in syntax. Doctoral dissertation, University of Massachusetts at Amherst.

Nishigauchi, T. 1990. *Quantification in the Theory of Grammar.* Kluwer.

Nunes, J. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation, University of Maryland.

Pesetsky, D. 1987. Wh-in-Situ: Movement and Unselective Binding. In E. Reuland and A. ter Meulen, eds., *The Representation of (In)definites,* MIT Press.

Saito, M. 1982. Case-marking in Japanese: A preliminary study. ms. MIT

Saito, M. 1992. Long distance scrambling in Japanese. *Journal of East Asian Languages,* 1.

Saito, M. 1998. Wh-Licensing in Japanese: Q-Wh Binding and Unselective Binding. A talk given at MIT.

Sohn, K. W. 1994. Adjunction to argument, free ride and a Minimalist program. MITWPL vol. 24. Department of Linguistics and Philosophy, MIT.

Tsai, W.-T. D. 1994. On Economizing the Theory of A-bar Dependencies. Doctoral dissertation, MIT.

Uriagereka, J. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26, 79-123.

Uriagereka, J. 1998. *Rhyme and Reason.* MIT Press, Cambridge.

Uriagereka, J. 1999. Ling 819: Minimalist Issues. An online seminar. University of Maryland.

Watanabe, A. 1992. Subjacency and s-structure movement of wh-in situ. *Journal of East Asian Linguistics 1,* 255-91.

Watanabe, A. 1995. Conceptual basis of cyclicity. MITWPL vol. 27. Department of Linguistics and Philosophy, MIT.

Yoshimura, N. 1992. Scrambling and anaphora in Japanese. Doctoral dissertation. USC. Los Angeles: CA.

Wh-Category Movement and the Legibility Problem of the Human Language Faculty

Koji ARIKAWA

The human faculty of language (FL) is legible to the faculty of performance (FP = articulatory-perceptual (AP) ╱sensorimotor systems + conceptual-intentional (CI) systems). How good a solution is FL to the legibility conditions that are imposed by FP (the *Legibility Problem*)? The *Linear Correspondence Axiom* (LCA) is a legibility condition imposed at the FL-AP interface (PF). The principle of *Full Interpretation* (FI) is a legibility condition imposed at the FL-CI interface (LF). I propose that FL's solution to the LCA and to the FI is optimal in the sense that it respects the Economy conditions, e.g., the minimization of derivational steps, the lack of superfluous steps, and the last resort nature of operations. At the descriptive level, I show that English, Hindi, and Japanese exhibit the identical computational procedure with respect to wh-movement. The multiple-wh effect in Japanese can be handled by the *Extended (Cyclicity) Condition* alone, which is a realization of negentropy within the human FL. Given the *Overt Wh-Category Movement Hypothesis*, the argument╱adjunct-wh asymmetry with respect to island effects is accounted for in a simpler and more natural way. The standard government-based ECP account is dispensed with.