Modal Feature Erasure and the Principle Last Resort

—Evidence for Another Formal Feature of $C^\circ$—

Koji ARIKAWA

1. Introduction
1.1. The Blocking Effect of Feature of $C^\circ$

A [+WH] feature in $C^\circ$ has a blocking effect on movement. Consider the following examples.

(1)a. *How1 do you wonder [which problem2 John could solve t2 t1]?
    $C1$ [+WH] $C2$ [+WH]

b. How1 do you think [(that)John could solve the problem t1]?
    $C1$ [+WH] $C2$ [−WH]

(Rizzi 1990: 8–9)

In (1a), the movement of an adjunct wh-phrase how1 is blocked by $C2$ in the lower clause which checks the WH-feature of an argument wh-phrase which problem2. One account for (1a) is as follows. When the derivation has reached [$C1$ you wonder [which problem John could solve t2 how1]]], the matrix $C1$ must attract the closest wh-phrase,
i.e., which problem2. However, a non-closest wh-phrase how1 is moved to C1. The derivation crashes due to the Minimal Link Condition, which requires the closest element be attracted (Chomsky 1995: 311, Kitahara 1997: 63). In (1b), C2 with a [−WH] feature does not block the adjunct-wh movement.

We propose another feature (a sublabel F’) of C0, which exhibits a similar blocking effect in Japanese sentences containing causal conjunctions. Since causal conjunction structures have no connection with the [±WH] feature, we argue for a non-WH feature residing in C0 which plays a role at LF.

1.2. Modal Asymmetry of Causal Conjunction Structures

There are various forms of causal conjunction (CC) in Japanese that are roughly translated into English as “so” or “because,” but which also feature different nuances. The following examples demonstrate that causal-conjunction clauses exhibit asymmetrical behavior in terms of the choice of the matrix-clause modality, e.g., some CC structures permit an imperative modal on the matrix clause, whereas some CC structures do not.

(2)

a. ame-ga hutte-ru-kara uti-de asobinasai
   rain-nom falling-nonpast-KARA home-at play-imperative
   ‘It is evident that it is raining, so play inside.’
Modal Feature Erasure and the Principle Last Resort

b. ame-ga hutte-ru-karakoso uti-de asob-inasai
   rain-nom falling-nonpast-KARAKOSO home-at play-imperative
   ‘(lit.) It is raining, and this is the very reason that I tell you to play inside.’

c. ame-ga hutte-ru-karaniwa uti-de asob-inasai
   rain-nom falling-nonpast-KARANIWA home-at play-imperative
   ‘(lit.) Now it is a decisive factor that it is raining, I naturally tell you to play inside.’

d. ??ame-ga hutte-ru-node uti-de asob-inasai
   rain-nom falling-nonpast-NODE home-at play-imperative
   ‘It is raining, therefore play inside.’

e. *ame-ga hutte-ru-okagede uti-de asob-inasai
   rain-nom falling-nonpast-OKAGEDE home-at play-imperative
   ‘(lit.) Thanks to the fact that it’s raining, play inside.’

f. *ame-ga hutte-ru-seide uti-de asob-inasai
   rain-nom falling-nonpast-SEIDE home-at play-imperative
   ‘(lit.) As a consequence of it raining, I tell you to play inside.’

g. *ame-ga hutte-ru-dakeni uti-de asob-inasai
   rain-nom falling-nonpast-DAKENI home-at play-imperative
   ‘(lit.) As might be expected of raining, I tell you to play inside.’
h. *ame-ga hutte-ru-bakarini uti-de asob-inasai
   rain-nom falling-nonpast-BAKARINI home-at play-imperative
   ‘(lit.) To my regret, since it is raining (and I wish the reverse),
   play inside.’

i. *ame-ga hutte-ru-tameni uti-de asob-inasai
   rain-nom falling-nonpast-TAMENI home-at play-imperative
   ‘(lit.) Raining is responsible for my telling you to play inside.’

j. *ame-ga hutte-ru-gayueni uti-de asob-inasai
   rain-nom falling-nonpast-GAYUENI home-at play-imperative
   ‘(lit.) It is raining, so it naturally follows that I tell you to play in-
   side.’

k. *ame-ga hutte-ru-toatte uti-de asob-inasai
   rain-nom falling-nonpast-TOATTE home-at play-imperative
   ‘(lit.) Raining may be just one of the reason that I tell you to play
   inside.’

l. *ame-ga hutte-ru-monodakara uti-de asob-inasai
   rain-nom falling-nonpast-MONODAKARA home-at play-imperative
   ‘(lit.) It is raining (which is an unexpected event), so play in-
   side.’

m. *ame-ga hut-te uti-de asob-inasai
   rain-nom falling-nonpast-TE home-at play-imperative
   ‘(lit.) It’s raining, and I tell you to play inside.’
n. *ame-ga hutte-ru-gen’ in-wa uti-de asob-inasai
   rain-nom falling-nonpast-cause-top home-at play-imperative
   ‘(lit.) The cause of raining is that I tell you to play inside.’

o. *ame-ga hutte-ru-riyuu-wa uti-de asob-inasai
   rain-nom falling-nonpast-reason-top home-at play-imperative
   ‘(lit.) The reason why it is raining is that I tell you to play inside.’

Now let us test how these CC structures interact with other matrix modal types. The following chart summarizes the restriction on the matrix clause modality with the variety of causal clauses in Japanese (cf. Teramura 1982, 1984). We represent a causal conjunction structure as P-CC Q, where P=causal clause, CC=causal conjunction, and Q=matrix clause.
<table>
<thead>
<tr>
<th>Modal of CC</th>
<th>Declarative</th>
<th>Conclusive</th>
<th>Assumptive (judgment)</th>
<th>Assumptive (explanation)</th>
<th>Persuasive</th>
<th>Inferential-Intentional (yoo)</th>
<th>Invitation</th>
<th>Imperative</th>
<th>Question</th>
<th>Question (explanation)</th>
<th>Judgment of Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>ru/ta</td>
<td>ru/ta</td>
<td>daroo</td>
<td>nodaro</td>
<td>noda</td>
<td>no</td>
<td>nemanba</td>
<td>nasaj</td>
<td>noka</td>
<td>karada</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pkara</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pkara koso</td>
<td>?</td>
<td>?</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pkara niwa</td>
<td>x</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pokage de</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>O</td>
<td>x</td>
</tr>
<tr>
<td>Pseide</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pdaken</td>
<td>O</td>
<td>O</td>
<td>?</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pbkarini</td>
<td>O</td>
<td>O</td>
<td>?</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Ptameni</td>
<td>O</td>
<td>O</td>
<td>?</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pgayueni</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>O</td>
<td>x</td>
</tr>
<tr>
<td>Ptoatte</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>O</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>O</td>
<td>x</td>
</tr>
<tr>
<td>Pmono dakara</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pte</td>
<td>O</td>
<td>O</td>
<td>?</td>
<td>O</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pgenin wa</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Priyuu wa</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Chart 1.

We propose the following four types of causal conjunction (CC) in Japanese.

Type I: KARA, KARAKOSO, KARANIWA
Type II: NODE
Type III: TE, OKAGEDE, SEIDE, DAKENI, BAKARINI, TAMENI, GAYUENI, TOATTE, MONODAKARA
Type IV: GEN’IN-WA, RIYUU-WA

The following questions are posed: Why is it that Type I CCs show insensitivity to the nature of modality of the matrix clause? Why are
Modal Feature Erasure and the Principle Last Resort

Type II and Type III CCs sensitive to the matrix clause modality? Why is it that sentences with the matrix modal forms accompanied by an element no ‘that’ (e.g., nodaroo, nod, noka) are insensitive to the matrix modal, whereas the matrix modal forms without the element no (e.g., daroo, da, ka) induce the sensitivity to the matrix modality? Since modality is related to a syntactic head $C^0$, it is plausible to assume that some feature of $C^0$ is relevant to the observed asymmetry.

Section 2 introduces some theoretical background adopted here. In particular, we will adopt the Feature-based theory in the Minimalist Program (Chomsky 1995, Kitahara 1997, Collins 1997), which assumes that the syntactic-feature movement plays a role in syntax both overt, i.e., a pre-Spell-Out level and covert, i.e., a post-Spell-Out level (LF).

In section 3, we give possible answers to the questions posed above, and show how the asymmetry among various types of causal conjunction phrases (CCP) are systematically accounted for within the theory.

In section 4, we recapitulate our observations, and reconsider the status of modal feature in the Feature-based theory.

2. Move-F Theory

Consider a DP [a book], a determiner phrase headed by determiner a. This DP is introduced into the derivation with various features, as indicated in (3) (cf. Chomsky 1995: 277).

(3)a. [+D] (a categorial feature)
   b. [+3 person, +singular] (ϕ-features)
   c. [+Case α] (a Case feature)
   d. strong F (feature), where F is categorial
Categorial features (3a) and $\phi$-features of the DP (3b) make contribution to meaning, whereas Case features of the DP (3c) and other strong features (3d) make no contribution to meaning. The former types are called interpretable features (F [+Interpretable]), while the latter types are called uninterpretable features (F [−Interpretable]). Let us first consider categorial features and $\phi$-features of the DP. An expression a book is not a verb but a noun headed by a determiner. Thus, a book cannot be used when one wants to reserve rooms (e.g., I will book the room vs. *I will a book the room). The expression a book also has features such as third person and singular. The expression cannot be used when one has more than one book (e.g., I have three books vs. *I have three a book). Categorial features and $\phi$-features of a DP thus make contribution to meaning; they are F [+Interpretable]. Now consider Case features (3c). Case features of a DP can be distinct although the relevant semantic role involved is unique (e.g., They expect [she will win] vs. They expect [her to win]). Thus, the Case features of the italicized subject of the embedded clause can be either nominative or objective (accusative) although both subjects are assigned the identical semantic role, i.e., the AGENT of win. It follows that Case features of a DP make no contribution to meaning, i.e., F [−Interpretable]. Affixal features (attracting heads), the Case assigning features of T and V, and $\phi$-features of verbs and adjectives are features of a target K (e.g., The perfective auxiliary verb has in she has gone contains $\phi$-features [+3rd person, singular] and these features make no contribution to meaning). These formal features are called sublabels or F-bar and they are always [−Interpretable] (Chomsky 1995: 278). A strong [−Interpretable] feature must be eliminated before Spell-Out and thus has various morphological effects. A weak [−Interpretable]
Modal Feature Erasure and the Principle Last Resort

features must be erased at LF. All [-Interpretable] features must be eliminated before FI (Full Interpretation) applies at LF; otherwise, there is no way for \( C_{HL} \) (the computational system of human language) to determine the identity of [-Interpretable] features at LF since they make no contribution to meaning. A [-Interpretable] feature left at LF must interact with the particular functional head under the particular configuration. This necessity of interaction with the functional head is the syntactic motivation for computational principles such as Last Resort (LR) (i.e., Greed or Attract) and Minimal Link Condition (MLC) of \( C_{HL} \), which roughly state that a [-Interpretable] feature must move to the closest c-commanding functional head (see Chomsky 1995: 280, Kitahara 1997: 14)\(^1\). The definition of the principle Last Resort is given below.

(4) Principle Last Resort (LR)

\( H(\alpha) \) attracts \( \alpha \) only if \( \alpha \) enters into a checking relation with a sublabel of \( K \), where a sublabel of \( K \) is a feature of the zero-level projection \( H(K)^{\text{Onax}} \) (including those features adjoined to \( H(K) \)).

Assume \( H(K) \) to be a head of a functional category \( K \), and \( \alpha \) a feature in a category which is c-commanded by \( K \). A checking relation is understood to be an identity-confirmation procedure, i.e., in this case, the identity of a feature of \( \alpha \) is confirmed by an adjunction (=attraction) of \( \alpha \) to the head of the closest c-commanding functional category \( K \) that bears a feature (a sublabel \( F' \)) that is compatible with \( \alpha \). LR states that a functional head \( H(K)^{\text{Onax}} \) with a [-Interpretable] feature attracts a [-Interpretable] feature \( \alpha \) (Attract), and \( \alpha \) adjoins to \( H(K)^{\text{Onax}} \) by an operation Move. This operation takes place only if there is a need to
confirm the identity of \( \alpha \) (Greed). A [−Interpretable] \( \alpha \) is erased when checked. If there is no need, the attraction must not take place\(^2\). The MLC is defined as follows in (4) (Chomsky 1995: 311).

(5) **Minimal Link Condition (MLC)**

\[ H(K) \] attracts \( \alpha \) only if there is no \( \beta \), \( \beta \) closer to \( H(K) \) than \( \alpha \), such that \( H(K) \) attracts \( \beta \).

MLC states that a head of a functional category attracts the closest feature that can potentially enter into a checking relation only if it needs to\(^3\).

Case features are strong [−Interpretable] features. A formal feature like Case is obligatorily driven to interact with a functional head before Spell-Out. Put differently, by itself a Case feature is not interpretable to LF interpretation system in \( C_{HL} \) until it actually interacts with a functional head before Spell-Out. All [−Interpretable] features must be eliminated (erased) for convergence \(^4\). Since a derivation cannot proceed with any unidentified feature, \( C_{HL} \) seeks the solution for identifying the nature of Case \(^5\). The solution (=the last resort) is such that the Case feature interacts with the closest c-commanding functional head. The solution must be executed within the overt Syntax and must not be procrastinated. A Case feature of a DP must be checked against the Case feature of the relevant functional head and then erased before these features flow into LF. The interaction is accomplished by an operation Move-F (F=feature), which adjoins a Case feature of a DP to the closest c-commanding functional head that can potentially check the feature identity \(^6\).

The major data that have been used to develop the Move-F Theory
Modal Feature Erasure and the Principle Last Resort

are such Case features, which are checked and erased by moving to functional categories such as v (or AGRo) and T (or AGRs). Regarding to the functional category C^0, the main data involve WH features. A recent finding states that semantic features like θ-roles are weak formal features that do not have to be checked before Spell-Out, but that are [−Interpretable] in that they must be checked and erased at LF (Saito and Hoshi 1998). We add one more weak [−Interpretable] feature F, namely, a modal feature of T^0 that interacts with the functional category C^0, thereby offering more evidence for the Move-F Theory.

3. Modal Feature Erasure

3.1. KARA-Clause (Type-I CCP: the embedded clause is CP)

3.1.1. The Matrix Modal Restriction Asymmetry

Consider the following examples (CC=causal conjunction).

(6)a. ame-ga hut-te-i-ru-kara uti-de ason-de-ru
    rain-nom-falling-CC house-at playing-declarative
    ‘It is raining, so (they) are playing inside.’

    b. ame-ga hut-te-i-ru-kara uti-de asob-oo
    rain-nom-falling-CC house-at play-inferential-intentional
    ‘It is raining, so let’s play inside.’

    c. ame-ga hut-te-i-ru-kara uti-de asob-inasai
    rain-nom-falling-CC house-at play-imperative
    ‘It is raining, so play inside.’
For KARA-clause, no restriction on the choice of mood (modal) of the matrix clause is observed; it may be declarative (6a), inferential-intentional (6b), or imperative (6c). To the contrary, as a causal conjunction, TE exhibits a restriction on the modal choice for the matrix clause.

(7) a. ame-ga hut-te-i-te uti-de ason-de-ru
       rain-nom-falling-CC house-at playing-declarative
              ‘It is raining, so they are playing inside.’

    b. *ame-ga hut-te-i-te uti-de asob-oo
       rain-nom-falling-CC house-at play-inferential-intentional
              ‘It is raining, so let’s play inside.’

    c. *ame-ga hut-te-i-te uti-de asob-inasai
       rain-nom-falling-CC house-at play-imperative
              ‘It is raining, so play inside.’

Before we go into the syntactic mechanism responsible for the asymmetry, let us first clarify the nature of T₀ and C₀ in the embedded clause, i.e., the causal clause.

3.1.2. The Indeterminacy Nature of T₀

To see the self-indeterminate nature of T₀, consider the various modal meanings of /ta/. It has more than the simple past meaning.
(8) a. kinoo i-ta
    yesterday be-past-declarative
    'I was there yesterday.'

    b. ook-iku nat-ta
    big-become-have-declarative
    'You have grown.'

    c. asita ki-ta hito-kara hajimete kudasai
    tomorrow come-perfective person-from begin-please
    'Please start with the person who will come here tomorrow.'

    d. asita it-ta hoo-ga ii
    tomorrow go-subjunctive side-nom good
    'You had better go there tomorrow.'

    e. mosi kite kure-nakat-ta-ra tasukar-anakat-ta
    if come give-not-subjunctive saved-not-subjunctive
    'If you hadn't come, I would not have been saved.'

    f. a, at-ta!
    Oh be-perfective-expectation realization
    'Oh, here it is.'

    g. saa, kat-ta, kat-ta!
    come on buy-command buy-command
    'Come on, buy it.'
h. a, simat-ta! kyoo-wa kaigi-ga aru-n-dat-ta.
 Shoot! today-top conference-nom take place-that-be-perfective-recall
 ‘Shoot! I forgot I had a meeting today.’

Like a Case feature of D₀, the meaning of /ta/ (T₀) is dependent on
the higher functional head (C₀) in the sentence and cannot be deter-
mined by itself. /ta/ can function as a simple past (8a), perfective
aspect and present tense (8b), perfective aspect with respect to the ma-
trix event (8c), specific advice (8d), counterfactual event (8e), realiza-
tion of an expectation (8f), mild command (8g), or recall (8h). The
following examples show that /ru/ has more than the simple present.

(9)a. ima i-ru

now be-present
‘Someone is there now.’

b. asita-wa i-ru

tomorrow-top be-future
‘Tomorrow I will be there.’

c. kinoo [[raisuu soko-e ik-u] hito] -ni age-ta

yesterday next week there-to go-imperfective person-to give-past
‘yesterday I gave in to person who is going tomorrow.’

d. saikin jogingu-o si-te-i-ru

recently jogging-acc doing-habitual
‘I make it a rule to jog.’
Modal Feature Erasure and the Principle Last Resort

e. tetu-wa mizu-ni sizum-u
   iron-top water-in sink-propositional
   'Iron sinks in water.'

/ru/ can mean present (9a), future (9b), imperfective (9c), habitual (9d), or propositional (9e).

Let us assume that \( T^o \) is introduced into the derivation as a mor-
pheme either /ru/or/ta, and a sentence must choose either one
when the sentence is declarative and has any relevance to a temporal
meaning (Teramura 1982, 1984). As shown above, tense (\( T^o \)) real-
ized as either /ru/or/ta/may bear a modal meaning. Turning to
the KARACLause, we see that not all modals can occur in the cause
clause. Consider the following examples.

(10) a. soko-e ik-u kara...
   there-to go-nonpast-declarative so
   'I will definitely go there, so...'

b. soko-e it-ta kara...
   there-to go-past-declarative so
   'I surely went there, so...'

c. soko-e ik-u-daroo kara...
   there-to go-nonpast-probably so
   'He will probably go there, so...'

-69-
d. *soko-e ik-oō kara...
   there-to go-inferential-intentional so
   'I will go there, so...'

e. *soko-e ik-anaika kara...
   there-to go-invitation so
   'Why don’t you go there, so...'

f. *soko-e ik-e kara...
   there-to go-imperative so
   'Go there, so...'

g. *soko-e ik-u ka kara...
   there-to go-question so
   'Will you go there, so...'

We propose that the modal feature involved in KARA-clause in (10a), (10b), and (10c) is declarative-mood-compatible, the feature of $C^0$ which we indicate as $[+DEC]$. Thus, the modal features responsible for moods like declarative and assumption are grouped as $[+DEC]$. $[-DEC]$ is a modal feature which is responsible for moods like inferential-intentional, invitational, imperative, or interrogative. See Radford (1997: 286) for postulating an abstract declarative affix $Dec$, bearing a weak feature. $Dec$ is the declarative counterpart of the question affix $Q$.

Given that $C^0$ is the locus of modals, it is plausible to assume that a modal feature of the embedded-clause $T^0$ must be associated with the closest c-commanding $C^0$.

We assume that a modal feature $[+DEC]$
Modal Feature Erasure and the Principle Last Resort

of the embedded-clause T° of a causal conjunction clause is [−Interpretatable] since the feature make no contribution to the cause-effect semantics. Rather, the [+DEC] feature of an embedded-clause T° is minimally required for the embedded clause to be qualified as a clause. Put in the terms of the Minimalist Program, the embedded-clause C° with [+DEC], being [−Interpretatable], must attract [+DEC] of the embedded-clause T° at LF, which is also [−Interpretatable]. This is parallel to the case where a [−Interpretatable] Q-feature ( [+WH] ) in C° attracts the closest [+WH] feature in D°. Let us summarize our hypothesis.

(11)a. T° in the cause clause bears a formal feature [+DEC].
   b. [+DEC] of T° is weak.
   c. [+DEC] of T° is [−Interpretatable].

Given this much, we will show below that the matrix modal restriction of causal conjunction structures follows from the Principle Last Resort and the MLC.

3.1.3. The Modal Feature Erasure Mechanism

At some point in the derivation of KARA-clause, the embedded C° with [+DEC] is introduced to merge with the embedded TP. [+DEC] in the embedded T° moves and adjoins to C°. [+DEC] in the embedded T° is checked against [+DEC] in the locally c-commanding C° (the MLC requires it) and [+DEC] in both T° and C° are erased at this point. Notice that principle LR (Last Resort) forces this [+DEC] -movement at LF; otherwise, a [−Interpretatable] feature
[+DEC] will survive and the derivation crashes. Like a Case feature of D⁰, the modal identity of T⁰ must be checked against the closest c-commanding C⁰ as soon as possible (the MLC). The derivation so far is represented as in Figure 1. CC is a zero projection of Causal Conjunction Phrase (CCP).

(12) [[ame-ga hutte [T⁰ ru] C⁰] kara]
    rain-nom falling-nonpast-declarative so
    'It is definitely raining so...'

```
         CCP
        /   \
       CP    CC
      /  \\
     TP   kara
    /
   vP
  /   \              T
 VP  v

Figure 1.
```

In Figure 1, T⁰ in the embedded clause is introduced with a feature [+DEC]. After the embedded-clause C⁰ with a feature [+DEC] is introduced, Move-F adjoining [+DEC] in T⁰ to the embedded C⁰ at LF, which is the closest functional head that can check the feature [+DEC] of T⁰. The identity of the embedded T⁰ is checked against the embed-
Modal Feature Erasure and the Principle Last Resort

ded C\textsuperscript{o} (no feature mismatch observed) and [+DEC] is erased immediately. The last resort property of [+DEC] movement is supported by the X-bar Compatibility Test advocated by Saito and Fukui (1996). The X-bar Compatibility Test states that if a language is head-initial, the rightward movement involves Merge since the movement forms the X-bar Compatible Structure. The leftward movement involves adjacency, which is triggered by feature-checking obeying the Principle Last Resort. The adjoined position becomes Spec after Spec-head agreement has established. If a language is head-final, the rightward movement is triggered by feature-checking obeying the Principle Last Resort, whereas the leftward movement is optional (Merge) since the movement forms the X-bar Compatible Structure. Given that Japanese is head-final, [+DEC] movement is a rightward movement and it must obey the Principle Last Resort because the movement does not constitute the X-bar Compatible Structure. (However, see Kayne (1994), which argues that all languages are head-initial.)

Suppose the derivation proceeds and has reached the following stage.

(13) \[ [[[\text{ame-ga hutte} [t\text{o} ru] C\text{o}] kara] uti-de asobi- [C\text{o}nasai]] \]

rain-nom falling-nonpast-declarative so home-at play-imperative

'It is raining, so play inside.'

Assume that the matrix C\text{o} nasai is introduced to merge with vP. The matrix C\text{o} bears [−DEC], which is a weak [+Interpretable] feature. [−DEC] does not have to be erased since the existence of [+Interpretable] features does not cause the derivation to crash. Now the overall derivation is as shown in Figure 2.
The \([-\text{Interpretable}]\) feature \([+\text{DEC}]\) is checked and erased. 
\([-\text{DEC}]\) is \([+\text{Interpretable}]\) feature, so its survival does not lead the derivation to crash.

Consider (6a), repeated here as (14). In this case, \(T^o\) is introduced in the matrix clause.

(14) \([\text{ame-ga hutte } [T^o \text{ ru } C^o \text{kara }] \text{ uti-de} \]
    rain-nom falling-nonpast-declarative so home-at

    asonde \([T^o \text{ ru } C^o]\)
    playing-nonpast-declarative
Modal Feature Erasure and the Principle Last Resort

'It's raining, so they are playing inside (I'm sure).'

The history of the derivation is as follows.

1. [+DEC] in the embedded $T^o$ is adjoined to the locally c-commanding $C^o$ with [ + DEC] by Move-F. This is forced by the principle LR (Last Resort) and the MLC (Minimal Link Condition).
2. [+DEC] in the embedded $T^o$ and [+DEC] in the embedded $C^o$ are checked and erased at LF.
3. [+DEC] in the matrix $T^o$ adjoined to the locally c-commanding matrix $C^o$ with [+DEC] by Move-F. This is forced by the principle LR and the MLC.
4. [+DEC] in the matrix $T^o$ and [+DEC] in the matrix $C^o$ are checked and erased at LF.
5. No [-Interpretable] feature survives, and the derivation may converge.

3.2. TE-Clause (Type III CCP: the embedded clause is TP)

Consider the following examples.

16a. ame-ga hut-te-i-te uti-de ason-de-i-ru
    rain-nom falling-CC house-at playing-declarative
    'It is raining, so (they) are playing inside.'

b. *ame-ga hut-te-i-te uti-de asob-roo
    rain-nom falling-CC house-at play-inferential-intentional
    'It is raining, so let's play inside.'
Examples of causal conjunction structure headed by TE feature a restriction on the choice of mood in the matrix clause, i.e., they allow the matrix C° to bear a declarative mood as in (16a), but not other moods such as inferential-intentional (16b) or imperative (16c). Suppose the derivation has reached the point as shown in the following.

(17) [ame-ga hut-te-i- [T° te]]
      rain-nom falling-declarative and
      ‘It’s raining and…’

We assume that there is no C° introduced into the derivation of TE-clause.

```
  TP
     \   /   \
      vP   T
        \   / \
         VP  v
         \       \  
        te [+DEC] 
```

Figure 3.

A causal conjunction clause headed by TE bears [+DEC], a minimum mood that is necessary for the clause to have a causal meaning. Since there is no closest c-commanding C° that can attract [+DEC] in the embedded T° at this point, the derivation proceeds without checking and
Modal Feature Erasure and the Principle Last Resort

erasure of \([+\text{DEC}]\).

Suppose the derivation has reached at the point indicated in Figure 4.

\[(\text{ame ga hut-te-i-} [T^0 \text{ te}]) \text{ uti-de ason-de-} [T^0 \text{ ru}] C^0 \]

rain-nom falling-declarative and home-at playing-nonpast-declarative

'It's raining, so they are playing inside.'

\[
\begin{array}{c}
\text{CP} \\
\text{TP} \\
\text{C [+] DEC} \\
\text{vP} \\
\text{T [+DEC] (\beta)} \\
\text{VP} \\
\text{v} \\
\text{TP} \\
\text{V} \\
\text{vP} \\
\text{T} \\
\text{te [+DEC] (\alpha)} \\
\end{array}
\]

Figure 4.

\([+\text{DEC}] (\beta)\) in the matrix \(T^0\) is closer to the matrix \(C^0\) than \([+\text{DEC}] (\alpha)\) in the embedded \(T^0\), because \(\beta\) c-commands \(\alpha\), and \(\beta\) is not in the minimal domain of the chain headed by an element which is adjoined to \(C^0\); in general, heads are excluded from the minimal domain of a chain. Hence, the MLC requires matrix \(C^0\) attract the closer \([+\text{DEC}]\) in the matrix \(T^0\) as in Figure 5.
Figure 5.

In the above structure, [+DEC] in the matrix $T^{o}$ is checked and erased. But [+DEC] in the matrix $C^{o}$ must be able to escape erasure for this first phase of the modal feature checking. That is to say, if [+DEC] of the matrix $C^{o}$ is erased, the derivation must crash since [+DEC] in the embedded $T^{o}$ will not be checked (a violation of Greed). On the contrary, the sentence in (18) is good; therefore there must be a way in which [+DEC] of the embedded $T^{o}$ can be checked successfully.

There is empirical evidence for the erasure-free nature of functional categories among languages. It is reported that in some languages, functional categories can never escape erasure of strong features when checked, while in other languages they can. It has been proposed that
the relevant parameter is the **Multiple Specifier Parameter** (see Chomsky 1995: 281, 286, 354, Kitahara 1997: 51, Koizumi 1995).

(19) **Multiple Specifier Parameter**

In some languages, a functional head is assigned the strong feature with an option to escape erasure when checked.

For example, $T^0$ in Icelandic and $T^0$ in Japanese are switched ON for this parameter with respect to syntactic Case-feature checking, whereas $T^0$ in English is switched OFF ($=EPP$ effect: all values of $T^0$, including infinitives, induce the EPP in English (Chomsky 1995: 282)). If a language has a positive setting for this parameter, a construction such as the multiple-subject construction (MSC) appears in the language (Jonas 1995, 1996).

(20) a. (Icelandic)

    pað hafa nokkrar kökur verið bakaðar fyrir veisluna.

    there have some cakes been baked for the-party

    ‘Some cakes have been baked for the party.’

b. (Japanese)

    kodomo-ga kotoba-ga yoku dekiru.

    children-nom language-nom good able

    ‘It is children who are skillful at languages.’

For the Icelandic example (20a), Chomsky argues that the functional head $T^0$ projects two specifiers for Case-feature checking: one for Expletive, and the other for Subject. In Japanese example (20b), the
functional head T^0 projects two specifiers for Case-feature checking; one for the first nominative DP, and the other for the second nominative DP.

Since we are dealing with weak features at LF, the parameter is too restrictive and it must be loosened. We propose the following as the modified version of the parameter.

(21) **Multiple Specifier Parameter (modified)**

In some languages, a functional head is assigned the [−Interpretable] feature with an option to escape erasure when checked.

That is to say, a week [−Interpretable] feature F called [+DEC] in Japanese C^0 has a parametric option such that it is deleted (not erased) for the first checking, and then it is erased for the second checking at LF. In other words, the [−Interpretable] feature [+DEC] in C^0 works twice for checking.

Given this much, it is reasonable to assume that Japanese C^0 is also switched ON for the Multiple Specifier Parameter (modified) with respect to Modal-feature checking. It follows that [+DEC] in the matrix C^0 is deleted (but not erased) when [+DEC] in the matrix T^0 adjoins to it for checking.

[+DEC] in the matrix T^0 is erased, but [+DEC] in the matrix C^0 is not. In the next stage of the derivation, [+DEC] in the embedded T^0 is moved by Move-F and is adjoined to the matrix C^0, a movement of which is driven by the principle LR and the MLC: being [−Interpretable], [+DEC] in the embedded T^0 cannot survive, and the matrix C^0 is the only closest C^0 locally c-commanding the [−Interpretable] feature [+DEC] of the lower T^0. The history of the derivation is shown
in Figure 6.

(22) [[ame ga hut-te-i- [T⁰ te]] uti-de ason-de- [T⁰ ru] C⁰]
    rain-nom falling-and home-at playing-nonpast
    'It's raining, so they are playing inside.'

Figure 6.

[+DEC] in the embedded T⁰ is forced to move to the matrix C⁰ by the
principle LR and the MLC. Thus, every [-Interpretable] feature is
checked and erased, thereby leading the derivation to converge.
Next consider the deviant example (16c), repeated here as (23).

(23) *ame-ga hut-te-i-te uti-de asob-inasai
    rain-nom falling-CC house-at play-imperative
    'It is raining, so play inside.'
Suppose the derivation has reached the following point.

(24) [ame-ga hut-te-i- [T° te]]
     rain-nom falling-declarative and
     ‘It’s raining, and…’

The embedded T° bears [+DEC]. Since there is no c-commanding C° that can check [+DEC] at this point, [+DEC] remains, and the derivation proceeds. Suppose the derivation has reached the following stage.

(25) * [[ame-ga hut-te-i- [T° te]] uti-de asobi-i- [C° nasai]]
     rain-nom falling-declarative and home-at play-imperative
     ‘It’s raining, and play inside.’

A [−Interpretable] feature [+DEC] in the matrix C° is the only c-commanding C° that can potentially check and delete [+DEC] in the embedded T°. The derivation is represented as in the following Figure 7.

(26) * [[ame-ga hut-te-i- [T° te]] uti-de asobi-i- [C° nasai]]
     rain-nom falling-declarative and home-at play-imperative
     ‘It’s raining, and play inside.’
Fig. 7.
Since a sublabel (feature) $F'$ of the matrix $C^o$ is weak ($[+\text{Interpretable}]$), the matrix $C^o$ does not attract $[+\text{DEC}]$ in the embedded $T^o$. Alternatively, $[+\text{DEC}]$ in the embedded $T^o$ may adjoin to the matrix $C^o$ by Greed. In that case, however, $[+\text{DEC}]$ can not be checked due to feature mismatch. As a consequence, $[-\text{Interpretable}]$ feature $[+\text{DEC}]$ of the embedded $T^o$ is not erased at LF and the derivation crashes. Thus, the ungrammatical status of the sentence in (26) parallels that of the following examples of checking failure.

(27) a. **Case-Feature Checking Failure**
    
    *John’s belief (of) [Mary to be intelligent]
    
    (cf. John believes [Mary to be intelligent])

b. **WH-Feature-Checking Failure**
    
    *dare-ga [Hanako-ga sono hon-o kat-ta-ka] sit-te-i-ru
    who-nom Hanako-nom that book-acc buy-past-Q know-nonpast
    'Who knows Hanako bought that book?' (Harada 1977)
Taroo-top who-nom that book-acc buy-past-Q know-nonpast
'Taroo knows who bought that book.'

In (27a), the Case feature of DP [Mary] is not checked, thereby violating Greed. In (27b), the WH-feature of the matrix subject DP [dare-ga] (who-nom) is not checked: the embedded \( C^0 \) bears \([+WH]\) feature, but it does not c-command the matrix subject \( D^0 \). Therefore, the embedded \( C^0 \) cannot attract \([+WH]\) in the matrix subject \( D^0 \) (the MLC violation). Thus, \([+WH]\) in the matrix subject \( D^0 \) is in violation of Greed. Likewise, the following example exhibits Modal Feature Checking Failure.

(28) **Modal Feature Checking Failure**

* [[ame-ga hut-te-i- [\( T^o \) te]] uti-de asobi- [\( C^0 \) nasai]]

rain-nom falling-declarative and home-at play-imperative

'It's raining, and play inside.'

The remaining existence of the \([-\text{Interpretable}]\) feature \([+\text{DEC}]\) of the embedded \( T^0 \) causes the derivation to crash \(^{10}\). For unchecked strong features, the cancellation of the derivation is guaranteed by the Strong Feature Condition defined as follows (SFC) (Kitahara 1997: 32).

(29) **Strong Feature Condition (SFC)**

Spell-Out applies to \( \Sigma \) only if \( \Sigma \) contains no category with a strong feature.
Modal Feature Erasure and the Principle Last Resort

The SFC basically states that $C_{hl}$ cannot produce an actual output if any strong feature survives in the final stage of derivation in the overt syntax. Spell-Out cannot apply to the derivation presented in Figure 7. Since our main concern is weak $[-\text{Interpretable}]$ feature that must be checked at LF, we propose the following Feature Condition that applies at LF.

(30) **Feature Interpretability Condition (FIC)**

FI (Full Interpretation) applies to $\Sigma$ only if $\Sigma$ contains no category with a $[-\text{Interpretable}]$ feature.

FIC states that $C_{hl}$ cannot produce an output if a $[-\text{Interpretable}]$ feature survives in the final stage of the derivation in LF.

One may argue against this analysis, however, since our account cannot explain the following examples:  

(31) a. *ame-ga hut-te-i-ru okage-de ie-de asob-inasai  
    rain-nom falling grace-by home-at play-imperative  
    ‘Thanks to it raining, play inside.’

b. ame-ga hut-te-i-ru okage-de ie-de asob-e-ta  
    rain-nom falling grace-by home-at play-can-past-declarative  
    ‘Thanks to it raining, we could play inside.’

c. *Taroo-no okage-de ie-de asob-inasai  
    Taroo-gen grace-by home-at play-imperative  
    ‘Thanks to Taroo, play inside.’
d. Taroo-no okage-de ie-de asobete

Taroo-gen grace-by home-at play-can-past-declarative

‘Thanks to Taroo, we could play inside.’

The modifiers of okage in (31c) and (31d) do not contain $T^o$ and therefore do not contain [+DEC], but they still exhibit the same contrast as that between (31a) and (31b), thereby showing that what is responsible for the contrast between (31a) and (31b) is not a feature like [+DEC]. More specifically, if [+DEC] is responsible for the contrast, our analysis will predict that (31c) must be acceptable, since there is no unchecked [+DEC] that will cause the derivation to crash.

However, it is too naive to assume that there is no $T^o$ in the modifier part of (31c) and (31d). The recent approach to possessives postulates IP which is a complement of D/P$^o$, a prepositional determiner (Kayne 1993/1994). Thus, the structure for John’s car is as follows.

\[(D/P^o)_{IP} \text{John } [\text{’s } \text{car}]\]

(Kayne 1994: 105)

In Kayne’s theory, ’s is $I^o$ and John occupies Spec of IP. If this is the right approach to possessives, the subordinate portion in (29c/d) will be as follows.

\[(D/P^o)_{TP} \text{Taroo } [\text{no } \text{okage}]\]

In (33), no is $T^o$ and one can hypothesize that it contains a feature [+DEC]. We can now maintain that (31c) is faulty since [+DEC] in the embedded $T^o$ is unchecked, whereas (31d) is proper because
Modal Feature Erasure and the Principle Last Resort

[+DEC] in the embedded $T^o$ is checked. Thus, the paradigm in (31) constitutes supporting evidence for our approach.

3.3. NODE-Clause (Type-II CCP: a borderline case of TP and CP)

Next, consider the following contrast among causal clauses headed by NODE.

(34)a. ame-ga hutteru-node uti-de ason-de-ru
rain-nom falling-CC house-at playing-declarative
'It is raining, so they are playing inside.'

b. ???ame-ga hutteru-node uti-de asob-oo
rain-nom falling-CC house-at play-inferential-intentional
'It is raining, so let’s play inside.'

c. ???ame-ga hutteru-node uti-de asob-inasai
rain-nom falling-CC house-at play-imperative
'It is raining, so play inside.'

We propose that the murky judgment of (34b) and (34c) is induced by the nature of the causal conjunction NODE, i.e., the CCP headed by NODE takes TP at its complement, but there is also an indication that CP is evolving.

Let us assume, for the sake of discussion, that NODE-clause is TP. The murkiness is then attributed to the incomplete nature of the evolving CP. The derivation of (34a) proceeds as follows.
(35) ①[ame-ga hut-te-i- [T_0 ru] node]
        [+DEC]

        [+DEC] in the embedded T_0 survives, since there is no functional head C_0 in NODE-clause.

②[[ame-ga hut-te-i- [T_0 ru] node] uti-de ason-de [T_0 ru] C_0]
        [+DEC]  [+DEC]  [+DEC]

        [+DEC] in the matrix T_0 is checked against [+DEC] in the matrix C_0 and is deleted (not erased). The deletion is represented by dotted line. [+DEC] in the matrix C_0 survives since the Multiple Specifier Parameter is switched ON for C_0 in Japanese.

③[[ame-ga hut-te-i- [T_0 ru] node] uti-de ason-de [T_0 ru] C_0]
        [+DEC]  [+DEC]  [+DEC]

The matrix C_0 attracts [+DEC] in the embedded T_0, which is erased by checking. [+DEC] in the matrix C_0 is erased this time. Every [−Interpretable] feature is checked and erased. The derivation converges.

The derivation of (34c) proceeds as follows.

(36) ①[ame-ga hut-te-i- [T_0 ru] node]
        [+DEC]

        [+DEC] in the embedded T_0 survives, since there is no C_0 in NODE-clause.

②[ [ame-ga hut-te-i- [T_0 ru] node] uti-de asob-i- [C_0 nasai]]
        [+DEC]  [−DEC]

−88−
Modal Feature Erasure and the Principle Last Resort

The matrix $C^\circ$ bearing a $[+\text{Interpretable}]$ feature does not attract $[+\text{DEC}]$ in the embedded $T^\circ$. The checking procedure for $[+\text{DEC}]$ fails (Greed violation). The $[-\text{Interpretable}]$ feature $[+\text{DEC}]$ in the embedded $T^\circ$ survives.

(3) FI fails to apply by FIC, and the derivation crashes.

However, if we assume that there is a CP, even if incomplete, $[+\text{DEC}]$ in the embedded $T^\circ$ in NODE-clause in (2) is “incompletely” checked off, i.e., not a total erasure of $[+\text{DEC}]$. Since the Multiple Spec Parameter does not affect $[+\text{DEC}]$ of $T^\circ$, the incomplete erasure is responsible for the marginal status of the sentence in question.

3.4. Type-IV CCP (Categorial Feature Checking within CP)

Consider the following contrast.

(37)a. *[[uti-de asonderu] riyuu]-wa ame-ga hutteru
    home-at playing reason-top rain-nom falling
    ‘The reason why (we) are playing inside is because it’s raining.’

b. [[uti-de asonderu] riyuu]-wa ame-ga hutteru karada
    home-at playing reason-top rain-nom falling because-be
    ‘The reason why (we) are playing inside is because it’s raining.’

We propose the following mechanism to account for the above contrast. Consider the structure for (37a).
In Figure 8, the categorial feature of 里由 "reason" $[+N, -V]$ is distinct from the feature of the matrix C, which is $[-N, +V]$. This causes feature-mismatch, and the derivation is canceled. Next, consider the structure for (37b).

In this case, the categorial feature of 里由 "reason" matches the feature of the matrix C. We propose that the feature-mismatch is avoided as long as there is a feature intersection as represented below.
Modal Feature Erasure and the Principle Last Resort

\[ (38) \quad \text{Feature Mismatch} \]
\[
\begin{array}{ccc}
N^0 & V^0 & \\
+N, -V & -N, +V & \\
\end{array}
\]

\[ \text{Feature Match} \]
\[
\begin{array}{ccc}
N^0 & P^0 & \\
+N, -V & -N & \\
\end{array}
\]

$N^0$ and $V^0$ have no feature intersection, whereas $N^0$ and $P^0$ do.

3.5. Further Evidence for Modal Feature Erasure

The following contrast in \((39)\) supports our hypothesis. The schematic representations are indicated in \((40)\).

\[ (39) \]

a. *ame-ga hutteru sei-de uti-de asonderu daroo
rain-nom falling consequence-be home-at playing-nonpast-assumption
'It's raining, and as a consequence, (they) are probably playing inside.'

b. *ame-ga hutteru sei-de uti-de asonderu ka?
rain-nom falling consequence-be home-at playing-nonpast-question
'It's raining, and as a consequence, are (they) playing inside?'

c. ame-ga hutteru sei-de uti-de asonderu no daroo
rain-nom falling consequence-be home-at playing-nonpast-that-assumption
'It's raining, and as a consequence, its the case that (they) are probably playing inside.'
d. ame-ga hutteru sei-de uti-de asonderu no ka?
   rain-nom falling consequence-be home-at playing-nonpast-question
   ‘It’s raining, and as a consequence, is it the case that (they) are
   playing inside?’

(40)

a. *[[...T° CC^{m}] …T° [C° daroo]]      (=39a)
b. *[[...T° CC^{m}] …T° [C° ka]]         (=39b)
c. [[...T° CC^{m}] …T° [C° no][C° daroo]] (=39c)
d. [[...T° CC^{m}] …T° [C° no][C° ka]]   (=39d)

CC^{m} indicates Type-III causal conjunction. We assume that C° like daroo and ka bear both [+DEC] (a [−Interpretable] feature) and [-DEC] (a [+Interpretable] feature). Crucially, such C° is switched OFF for Multiple Subject Parameter. Consider the history of the derivation for (39a).

---

Figure 10.
Modal Feature Erasure and the Principle Last Resort

A $[-\text{Interpretable}]$ feature $[+\text{DEC}]$ of the embedded $T^o$ survives, and the derivation crashes. Consider the history of the derivation for (39c).

![Diagram of the derivation]

**Figure 11.**

It is $[+\text{DEC}]$ of the functional head *no* that saves the sentence. $[+\text{DEC}]$ in the embedded $T^o$ can be checked by $[+\text{DEC}]$ in $C^o$ *no*. Every strong feature is checked and erased, thereby leading the derivation to converge. Notice that $[-\text{DEC}]$ is $[+\text{Interpretable}]$ and it need not be erased.

Notice that it is crucial to assume that $C^o$ with dual features is immune from the Multiple Specifier Parameter. That is to say, when $[+\text{DEC}]$ appears along with $[-\text{DEC}]$ in a head $C^o$, $[+\text{DEC}]$ must be erased when checked at the first checking, i.e., the Multiple Specifier Parameter is being switched OFF. Why is this the case? Our speculation is as follows. In a sense, this type of $C^o$ with dual features
displays multiple specifiers to begin with, i.e., the MultipleSpecifier Parameter is lexically ON from the beginning when it is introduced into the derivation. The MultipleSpecifier Parameter remains inactive for something that already exhibits multiple Specs.

4. Concluding Remarks

4.1 Types of Causal Conjunction and Modal Feature Erasure

Type-I CCs (causal conjunctions) have no restriction on the choice of mood of the matrix clause. A CC of this type merges with CP, and F (\([-\text{Interpretable}\) feature] \([+\text{DEC}\) of the lower \(T^0\) is checked and erased within this lower CP, thereby leaving no F before the derivation proceeds into the matrix level. At the matrix level, whatever F in \(T^0\) is checked and erased within the matrix CP, or there is no F at this level (as in the imperative). Thus, the computation leaves no F, leading the derivation to converge and yields a legitimate output by FIC.

Type-II CCs form a natural class in that they are etymologically related to nouns, e.g., OKAGE “the grace of God,” SEI “one’s doing things,” DAKE (\(<\text{TAKA}\) “all that there is,” BAKARI (\(<\text{HAKARI}\) “measure,” TAME “advantage,” YUE “reason,” and MONO “thing.” Thus, the hypothesis that the CCPs headed by Type-III CCs have a structure of a complex NP is plausible. There is evidence that a complex NP contains a TP, not a CP (Murasugi 1991). TE-clause must also be a TP by assumption. Thus, the Type-III CC merges with TP. A \([-\text{Interpretable}\) F \([+\text{DEC}\) in \(T^0\) will not be checked if there is no \([+\text{DEC}\) in the matrix \(C^0\), and the derivation crashes. This is the syntactic mechanism of the matrix modal restriction on this type of CC.
structure. The situation is similar in Type-II CC NODE. The marginal status of this CC is attributed to the "incomplete" erasure of [+DEC].

The behavior of type-IV CCs is accounted for by Spec-Head agreement with respect to the categorial feature checking.

4.2. The Status of Modal Feature in the Feature-based Theory

What is F, [−Interpretable] feature in natural language? There is evidence that semantic features as θ-roles are formal features that do enter into the syntactic feature checking at LF (Saito and Hoshi 1998)\textsuperscript{12}. There are cases where a syntactic principle like Relativized Minimality, which should be reduced to the MLC (Kitahara 1997: 61-65), is relevant to an operation that sees semantic features such as the Existential Closure as per Heim (1982)\textsuperscript{13}. These findings indicate that some semantic features are sensitive to syntactic principles of UG (Universal Grammar). If it is the case, it is not unreasonable to hypothesize that a modal feature like [+DEC] is also a formal feature that play a role in the syntactic feature checking.
Notes

* The basic idea of this article is presented at Nihongo kyooiku gakkai “A conference on Japanese education” held at Kyoto University of Foreign Studies in 1996. I would like to thank two anonymous reviewers of Nihongo Kagaku “The scientific approach to Japanese language” (Kokuritu kokugo kenkyuuzyo “National language research center”), who have critiqued the first draft of this article. As usual, there are agreements and disagreements among us. All the shortcomings contained here are mine. The research reported here was supported in part by Grant-in-Aid for Encouragement of Young Scientists from the Japanese Ministry of Education, Science, Sports and Culture (Project #: 10710260).

1) The definition of the principle Last Resort stated in Chomsky (1995: 280) is the following:

   Last Resort

   Move F raises F to target K only if F enters into a checking relation with a sublabel of K.

2) Deleted α is invisible at LF but still accessible to the computation, but erased α is inaccessible to any operation. Erase α is a stronger form of deletion (Chomsky (1995: 280))

3) The notion closer contains the notions of c-command (i) and equidistance (ii) defined as:

   (i) \( \alpha \) c-commands \( \beta \) iff every category dominating \( \alpha \) dominates \( \beta \), \( \alpha \neq \beta \), and neither dominates the other (Reinhart 1976).

   (ii) \( \beta \) does not prevent H(K) from attracting \( \alpha \) if \( \beta \) is in the minimal domain of CH, where CH is the chain headed by \( \gamma \), and \( \gamma \) is adjoined to H(K) (Chomsky 1995: 299).

The notion closer is now defined as:
(iii) $\beta$ is closer to $H(K)$ than $\alpha$ iff $\beta$ c-commands $\alpha$, and $\beta$ is not in the minimal domain of $CH$, where $CH$ is the chain headed by $\gamma$, and $\gamma$ is adjoined to $H(K)$ (Chomsky 1995: 299).

4) See Chomsky (1995: 277). He suggests that interpretability at LF relates only loosely to the intrinsic-optional distinction (Chomsky 1995: 278). Collins (1997: 139, fn. 4) suggests that the distinction between interpretable and uninterpretable features may be more adequately characterized as the distinction between inherent and relational features.

5) This constitutes a part of answers for the questions raised by Chomsky (1995: 278). The two questions are:

(1) Why is a sublabel $F'$ of the target that enters a checking relation invariably $[-\text{Interpretable}]$?

(2) Why is $F'$ present at all? (= Why does language have the operation Move?)

For question (2), in resemblance to an immune system, $C_{ntl}$ needs $F'$ (= an "antibody") to identify an unknown $F$ (= an "antigen") which invades the derivation through random Mergers: $F'$ attracts $F$ and makes $F$ "nonpoisonous" for $C_{ntl}$.

6) According to Chomsky, the $[+\text{WH}]$ feature of a wh-phrase is $[+\text{Interpretable}]$. However, if we consider the following Japanese data, the meaning of wh-phrase is not inherent, but relational. That is to say, $[+\text{WH}]$ does not contribute to meaning.

(i) nani-o katta no?
what-acc bought Q
'what did you buy?'
(ii) kimi-wa sono nani-o katta n da ne?
you-top that doohickey-acc bought that is right
‘You bought that doohickey, right?’

A wh-phrase can be either “what” or “it” (doohickey=something the name of which you have forgotten) according to the nature of the matrix C°. That is, if every wh-phrase bears a [+WH] feature, the [+WH] feature may be [−Interpretable]. What about features like [+TOP] or [+FOCUS] ? NP-wa bears [+TOP]. The meaning of NP-wa can be either thematic or contrastive, according to which head the feature [+TOP] adjoins. If it adjoins to C°, the meaning is thematic, if it adjoins to T° or v, contrastive meaning arises.

(i) [sono hon-wa John-ga katta C+ [+TOP]]
that book-top John-nom bought
‘Speaking of that book, John bought it.’

(ii) sono hon-wa [John-ga kat [TA ta] + [+TOP]] C
that book-top John-nom bought
‘John bought that book, at least.’

(iii) John-ga sono [hon-wa [V kat] + [+TOP]] [TA ta] C
John-nom that book-top bought
‘John bought that book, at least.’

Thus, [+TOP] must interact with functional heads in order to identify itself. [+TOP] itself makes no contribution to meaning (thematic vs. contrastive). Hence, [+TOP] is [−Interpretable]. The feature
Modal Feature Erasure and the Principle Last Resort

[±FOCUS], which triggers the scrambling effect, is dependent on the information of functional categories such as v (triggering a short scrambling=A-movement) and T° (triggering a long scrambling=A/A′-movement). [+FOCUS] is irrelevant to anaphor-binding possibility, which is sensitive to A/A′-distinction. Therefore, [+FOCUS] is also [−Interpretable].

7) Evidence for the existence of C° in KARA-clause comes from adjunct-WH licensing. Consider the following contrast.

(i) [[Hanako-ga naze sono hon-o katta] kara] Taroo-wa okotteru no?
Hanako-nom why the book-acc bought CC Taroo -top be angry Q
‘* Why1 is Taro angry at her because Mary bought the book t1?’

(ii) * [[Hanako-ga naze sono hon-o kat-] -te] Taroo-wa okotteru no?
Hanako-nom why the book-acc buy-CC Taroo -top be angry Q
‘* Why1 is Taro angry at her because Mary bought the book t1?’

(iii) *? [[Hanako-ga naze sono hon-o katta]node] Taroo-wa okotteru no?
Hanako-nom why the book-acc buy-CC Taroo -top be angry Q
‘* Why1 is Taro angry at her because Mary bought the book t1?’

(iv) * [[Hanako-ga naze katta] hon-o Taroo-wa hosigatteru no?
Hanako-nom why bought book-acc Taroo -top want Q
‘* Why1 does Taro want the book which Hanako bought t1?’

In all these examples, an adjunct WH-phrase naze “why” is intended to modify the embedded-clause verb kaw- “to buy”. Being adjunct, the [+WH] feature of naze must adjoin to C° within the embedded clause
to be licensed. If the WH feature of naZe has to move and adjoin to the matrix C^o, it always crosses a barrier, namely, the embedded-matrix boundary. Only KARA-clause contains C^o, and the adjunct WH is licensed by it.

8) The head-adjointed position is contained in the checking domain of the head, following the definition of checking domain proposed in Chomsky (1995: 177-178).

The **checking domain** of α is the minimal domain of α minus the minimal complement domain of α.

The **minimal domain** of α is the smallest subset of K of the domain of α such that for any γ ∈ the domain of α, some β ∈ K reflexively dominates γ.

The **domain** of α is the set of nodes contained in MAX (α) that are distinct from and do not contain α.

The **complement domain** of α is the subset of the domain of α reflexively dominated by the complement of the construction.


10) The Principle Last Resort has two versions; the weaker version, i.e., Greed, and the strongest version, i.e., Greediest. The definition of Greed in Chomsky (1993, 1994: 14) and Chomsky and Lasnik (1993: 564) follows:

**Greed**

Move α raises to a position β only if morphological properties
Modal Feature Erasure and the Principle Last Resort

of \( \alpha \) itself would not otherwise be satisfied in the derivation.
The definition of Greediest in Collins (1997: 96) is as follows.

**Greediest**

Move \( \alpha \) raises to a position \( \beta \) only if some morphological property of \( \alpha \) itself is satisfied in position \( \beta \).

If we adopt Greed, \([+DEC]\) in \( \mathcal{T} \) does not move to \( \mathcal{C} \) since \([+DEC]\) is just one of the formal features of \( \mathcal{T} \). In this case, unchecked \([+DEC]\) causes the derivation to crash. If we adopt Greediest, \([+DEC]\) in \( \mathcal{T} \) moves to \( \mathcal{C} \) since \([+DEC]\) is sufficient for triggering feature movement. In this case, a feature mismatch causes the derivation to crash.

11) An anonymous reviewer of *Nihongo Kagaku* “Scientific Approach to Japanese Language,” (Kokuritsu Kokugo Kenkyuuzyyo (National Language Research Center) pointed out these as counterexamples. However, not only they do not constitute counterexamples, they actually support our approach.

12) Saito and Hoshi (1998) argues that an LF incorporation of a theta-assigning noun to a light verb takes for theta-role assignment and it is motivated by the Principle Last Resort. It follows that theta-roles are formal features.

13) Consider the second rule of Quantifier for selection indices (Heim 1982: 146).

(i) Copy the referential index of every indefinite NP as a selection index onto the lowest-c-commanding quantifier.

Heim explains that “the lowest c-commanding quantifier” is “one among all the quantifiers that c-command the indefinite which does not c-command any other one.” In our terms, it is the closest c-commanding quantifier. Consider the following example.

---
(ii) If a man is lonely, he often buys a cat.

In (ii), *often* is an adverb of quantification. The indefinite DP *a man* receives a universal interpretation, but the indefinite DP *a cat* does not. (ii) has the following structure.

(iii) $\text{[often ... a man $[\exists s \ldots \text{a cat}]]$}$

In (iii), $\exists$ (Existential quantifier) is the minimal quantifier for *cat*. Therefore, *cat* cannot have a universal reading.

References

Unpublished manuscript, University of Connecticut and Tohoku University.


Modal Feature Erasure and the Principle Last Resort


in Linguistics 27, 143-163.
Modal Feature Erasure and the Principle Last Resort
—Evidence for Another Formal Feature of $C^0$—

Koji Arikawa

The feature checking theory concentrates on categorial and syntactic features (cf. Chomsky 1995). It is assumed that the remaining existence of a $[-\text{Interactable}]$ feature at LF causes the derivation to crash. Move is forced to eliminate $[-\text{Interactable}]$ features. Structural Case features and some categorial features are $[-\text{Interactable}]$ and must be eliminated. A $[+\text{WH}]$ feature of $C^0$ and $D^0$ has long been regarded as a syntactic feature constrained by the Principle LR (Last Resort).

What about semantic properties? The current analysis of theta role discharge contends that theta roles are formal features of $V^o$ and that they obey the Principle LR (Saito and Hoshi 1998, Bošković and Takahashi 1995, Lasnik 1995).

We argue for the existence of another formal feature of $T^o$ and $C^o$, i.e., a modal feature. Modal feature checking is another LR-constrained UG operation, just as Case feature checking and WH feature checking.