CO-OCCURRENCE RESTRICTIONS ON ARGUMENTS IN THE THETA SYSTEM

KENJI SUGIMOTO

Tohoku University*

Keywords: theta system, co-occurrence restrictions, cluster distinctness, repair strategies, C-I interface

1. Introduction

The theta system proposed by Reinhart (2002) accounts for different syntactic realizations of the same lexical verb entry. However, this system cannot completely capture the co-occurrence restrictions on arguments in causative psych verbs. This paper attempts to provide a solution to this problem by deriving these restrictions from a general property of the C-I (conceptual-intentional) system and a repair mechanism applied at the C-I interface, while at the same time trying to maintain the advantages of the theta system.

This paper is organized as follows: in section 2, we will review the framework of the theta system proposed in Reinhart (2002) and point out its problems; in section 3, we will provide our proposals for solving the problems and further arguments in their favor; in section 4, we will present concluding remarks.

* This paper is a revised version of a paper presented at the 83rd General Meeting of the English Literary Society of Japan, held at the University of Kitakyushu on May 21–22, 2011. I would like to thank the participants to the meeting for their valuable comments. I would also like to express my sincere gratitude to Professor Yoshiaki Kaneko and Professor Etsuro Shima for their insightful comments and suggestions. I am also grateful to Yoshihito Dobashi, Cornelia Daniela Lupsa, Tadao Maruta, Nobuhiro Miyoshi, Taichi Nakamura, Takahiro Tozawa, and three anonymous reviewers for their helpful comments and constructive criticism. My thanks also go to the members of the Department of English Linguistics at Tohoku University. I am grateful to Max Phillips Jr., who has kindly acted as an informant. Last but not least, I would like to thank Cornelia Daniela Lupsa and Max Phillips Jr. for suggesting several stylistic improvements. All remaining errors are, of course, my own.

This section provides a brief overview of the system in Reinhart (2002) and some motivations for adopting it. Although thematic role labels such as agent and theme have been used as if they were primitive terms, she argues that they are not primitives and are composed of two types of formal features with two values: the cause change feature (+c-feature) and the mental state feature (+m-feature). The former codes the causal relation in the event, and the latter codes some mental properties of the participant in the event.

The +c-feature indicates that the role is perceived as a sufficient condition for the event denoted by a verb (Shen (1985)). The −c-feature indicates that the role is identified as not being the cause of the event. If the c-feature is not represented, the causal status of the role is linguistically underdetermined, which does not imply that the role cannot be a cause. The +m-feature reflects the mental state of the participant that motivates the event denoted by the verb. It entails animacy. However, the −m-feature does not entail inanimacy.

A cluster does not need to be specified for all features, so that in principle, the two binary features define the following nine feature clusters.¹

(1) Theta feature cluster²
a. [+c+m] — agent
b. [+c−m] — instrument (…)
c. [−c+m] — experiencer
d. [−c−m] — theme / patient
e. [+c] — cause (Unspecified for / m); consistent with either (a) or (b).
f. [+m] — sentient (?)
g. [−m] — (Unspecified for / c): subject matter / locative source
h. [−c] — (Unspecified for / m): Internal roles like goal, benefactor typically dative (or PP).
i. [ ] — Arb(bitrary) (Reinhart (2002: 232, 285))

¹ Reinhart (2002) mentions that the correspondence of these clusters to the traditional θ-roles is not always one to one, and that many of these clusters have different interpretations, depending on some properties of the lexical semantics of verbs.
² / α stands for the feature (and value) α.
Reinhart (2002) offers empirical motivations for adopting this θ-feature system. Among these motivations is a problem about θ-selection. In terms of the traditional θ-roles, a causative psych verb like worry takes an agent and a cause such as an instrument or a natural force as its external argument, as shown in (2a, b).

(2)  
\begin{align*}
a. & \text{ The doctor}_\text{agent} \text{ worried Lucie}_\text{experiencer}. \\
b. & \text{ The news}_\text{cause} \text{ worried Max}_\text{experiencer}. \\
c. & [\text{Agent, Experiencer}] (\text{ii}) [\text{Cause, Experiencer}] \\
d. & ([+c], [-c+m])
\end{align*}

Thus, the verb must have the two entries in (2c), which select different external θ-roles. The problem with this polysemy approach is that it is not clear how multiple entries are related to each other. In contrast, Reinhart’s (2002) θ-feature system makes do with the unique entry in (2d), which contains the external [+c] cluster, because the agent and cause roles have the +c-feature in common. This cluster is specified only for the c-feature with the m-feature unspecified in the lexical entry. In her system, whether this cluster is finally interpreted as an agent or a cause is determined by the context of use after syntactic computations. This monosemy approach is possible only by assuming her θ-feature system.

Further motivation comes from an analysis of unaccusative verbs. It has been observed that many unaccusative verbs also have a transitive alternative, which is known as the causative-inchoative alternation. The standard approach to this alternation assumes that the transitive entry is derived from the basic unaccusative entry. On this analysis, unaccusative verbs are just listed in the lexicon. However, this view raises certain learnability problems. Unaccusative and unergative verbs have significantly different syntactic realizations, so it is crucial for children to determine which one-place verbs are unaccusative. However, it is unclear how this knowledge is acquired. In opposition to the standard approach, Reinhart (2002) assumes with Chierchia (1989) that the transitive entry is basic and the unaccusative entry is derived from it by a lexical reduction operation. Under this view, children need to know which transitive verbs allow the alternation. She proposes a solution to this problem by assuming that the reduction operates on the transitive verbs taking agent and cause as a subject (e.g. worry, open and break), but not on those selecting only agent as a subject (e.g. feed and eat). The former contain the [+c] role in their base entry, while the latter contain the [+c+m] role. Therefore, only verbs selecting the [+c] role are subject to the reduction operation. If we assumed that transitive verbs such as worry, open and break have two entries which select different external
θ-roles, it would be difficult to solve the learnability problem. Thus, the θ-feature system has the merit of solving the learnability problem by enabling the verbs to have unique entries.

So far, we have briefly reviewed the theta system and its motivations. However, it is not without its problems. Let us reconsider the lexical entry of the verb worry.

(3) a. Lucie_{experiencer} worried about the doctor_{subject-matter}.
   b. ([+c], [−c+m], [−m])

In addition to the θ-roles in (2a, b) the verb may allow the subject-matter toward which the mental state is directed, as shown in (3a). On the grounds of such argument realizations, Reinhart (2002) assumes that the verb has the argument structure in (3b). This θ-grid undergoes several procedures and generates the syntactic derivations in (2a, b) and (3a) according to her linking system. Note that there is another syntactic structure generated from the θ-grid of (3b), in which it is possible for the [+c] and [−m] arguments to be realized together. However, such a derivation is not permitted as illustrated in (4).

(4) *The doctor’s letter_{[+c]} worried Lucie about her health_{[−m]}.
   (Reinhart (2002: 261))

Thus, the argument structure in (3b) overgenerates.

Reinhart (2002) argues that this may reflect a broader restriction, a generalization that the same θ-role cannot be realized twice. To solve this problem, she proposes the following condition.

(5) CD (Cluster Distinctness)
   a. Two indistinct θ-clusters cannot be both realized on the same predicate.
   b. Distinctness: Two feature-clusters α, β, are distinct iff (i) they share at least one feature, and (ii) there is at least one feature or value which they do not share.
   (Reinhart (2002: 264))

At the stage of the lexicon-syntax mapping, CD imposes co-occurrence restrictions not only on the same θ-cluster, but also on two indistinct θ-clusters as defined in (5b). The intuition behind (5bi) is that identifying distinctness requires some shared basis for comparison. In addition, (5bii) implies that some unshared feature or value makes two clusters distinct. Under (5b), the [+c] and [−m] clusters are indistinct, because they share no feature as the basis for comparison and so do not satisfy (5bi). Thus, Reinhart (2002) argues, (5) appropriately rules out argument realizations such as (4).

Reinhart’s (2002) motivation for applying CD at the lexicon-syntax inter-
face comes from the contrast between the lexical causative in (4) and the analytic causative in (6).

(6) The doctor's letter \([c]^-\) made Lucie \([-c+m]^-\) worry about her health \([-m]^-\].

(Reinhart (2002: 261))

Logically, the relevant two roles are compatible, because the content intended in (4) can be expressed with a different structure as in (6). This seems to lead Reinhart (2002) to argue that this contrast should be accounted for from a purely grammatical perspective, but not from an interpretive perspective. Specifically, the grammaticality of (6) can be captured by assuming that the \([+c]^-\) cluster is introduced by make, and the \([-m]^-\) cluster by worry. In other words, since the two \(\theta\)-clusters are introduced by different predicates, (6) is not incorrectly ruled out by CD.

As discussed above, in Reinhart's (2002) analysis, the \([+c]^-\) cluster in the lexical entry is consistent with an agent role and a cause role. This interpretative distinction is finally made by context after syntactic computations. Then, the application of CD at the lexicon-syntax interface should prevent not only the \([+c]^-\) cluster with a cause interpretation like (4), but also the one with an agent interpretation from co-occurring with the \([-m]^-\) cluster, contrary to fact.

(7) ?Albert \([+c]^-\) worried John \([-c+m]^-\) about the adequacy of his insurance coverage \([-m]^-\].

(Bouchard (1995: 333))

In spite of the simultaneous realization of the relevant arguments, (7) is acceptable in a certain context where Albert, John's insurance agent, was constantly reminding John that his insurance coverage was inadequate (Bouchard (1995)). (7) contrasts with (4) in that Albert plays an agent role rather than just a cause role in the event. This contrast is not a marked phenomenon found with the verb worry alone. In fact, our informant agrees that the same holds true of other causative psych predicates such as relax and bother.

   b. Kaori relaxed John with a Japanese dinner.

\(^3\) Reinhart (2002) assumes that an agent role consists of causation and volition.

\(^4\) Reinhart (2002) notes that the verb interest also has the same property, as shown in (i).

(i) a. Lucie / the article interested Max.
   b. Lucie interested Max in linguistics.
   c. */?The article interested Max in linguistics.

(Reinhart (2002: 265))
(9)  
\begin{align*}
\text{a. } & \text{The news bothered Lucie about environment problems.} \\
\text{b. } & \text{The math teacher bothered Lucie about a math problem.}
\end{align*}

Here, one might argue that in principle, these roles could be distinguished in the lexicon by assuming two entries for the verb, that is, one including the [+c] cluster and the other including the [+c+m] cluster. However, lexical polysemy would bring back the problems of θ-selection and learnability, as discussed above. To maintain the advantages of the theta system, we should not assume multiple entries for a single verb.

3. A New Approach to Co-occurrence Restrictions

As discussed in section 2, the problem with the theta system lies in the application of CD at the lexicon-syntax interface. In addition, although Reinhart (2002) argues that whether the [+c] cluster actually means an agent or a cause depends on context, she does not offer a clear way of linking them. These problems lead us to entertain the following new ideas.

(10)  
\begin{align*}
\text{a. } & \text{CD is an interpretative condition on the LF representation.} \\
\text{b. } & \text{The [+c] cluster is supplied with the +m-feature by context at the C-I interface.}
\end{align*}

(10b) enables the [+c] cluster to be disambiguated at the C-I interface. Then, according to (10a), CD interprets the [+c+m] and [−m] clusters as distinct clusters which are legible to the C-I system. This is illustrated in (11).

(11)  
\begin{align*}
\text{Base entry} & \rightarrow \text{CS (computational system)} \rightarrow \text{C-I system} \\
[+c] & \rightarrow [+c] & \rightarrow [+c+m] \\
[-m] & \rightarrow [-m] & \rightarrow [-m]
\end{align*}

In the following sections, we recast our intuition in the shape of clear definitions and offer a solution to the problem pointed out in the previous section.

3.1. Interpretative CD and the Feature-Suppletive Operation at the C-I Interface

First, in conformity with (10a), we replace clause (a) of Reinhart’s (2002) CD in (5) with the following interpretative cluster distinctness (ICD).

(12)  
\begin{align*}
\text{ICD: } & \text{The C-I system does not distinguish two indistinct θ-clusters.}
\end{align*}

Among the theoretical motivations for ICD is inability to distinguish indis-
CO-OCCURRENCE RESTRICTIONS ON ARGUMENTS IN THE THETA SYSTEM

375
tinguishable (IDI) proposed by Reuland (2011). He assumes IDI to be a
general property that any computational system must have.⁵ He goes on
to show that the core cases of Condition B of the binding theory can be
derived from IDI. More specifically, the reason for the violation is the in-
ability of the CS to distinguish between occurrences of identical variables
in logical syntax at the C-I interface. Since ICD implies in an analogous
fashion that if there are indistinct θ-clusters in the LF representation, the
interpretive system does not interpret them appropriately, ICD may be re-
garded as a reflex of IDI in the C-I system.

We also modify clause (b) of Reinhart’s (2002) CD as follows.

(13) Distinctness: Two feature-clusters α, β are distinct iff α is specified
as [+f] and β as [−f], for some shared feature f.

This new distinctness is the same as Reinhart’s (2002) original one in (14)
in that it requires some shared feature as the basis for comparison.

(14) Two feature-clusters α, β, are distinct iff (i) they share at least
one feature, and (ii) there is at least one feature or value which
they do not share.

However, our distinctness differs from clause (ii) of her original one. In
clause (ii), an unshared feature or value makes two feature-clusters dis-
tinct. In contrast, our distinctness refers only to the value of the shared
feature in deciding whether two feature-clusters are distinct or not. Let us
consider the [−c−m] and [−c] clusters. These clusters share the c-feature
as the basis for comparison. Under Reinhart’s (2002) distinctness, the
unshared m-feature makes these clusters distinct. In contrast, our distinct-
ness regards them as indistinct, because the shared basis for comparison, the
c-feature, has the same value. In this paper, we assume the distinctness in
(13), not Reinhart’s (2002) original one.

Next, let us turn to (10b). The feature-suppletive operation at the
C-I interface may be a device predicted from the whole language system
rather than a device stipulated in order to account for the co-occurrence
facts. Some support comes from research on the syntax-morphology inter-
fase. Halle and Marantz (1993) propose the late insertion model of phono-
logical features to capture both the systematic relationships between syntax
and morphology and syntax-morphology mismatches. If the whole system

⁵ Among representatives may be the antilocality condition on movement in syntax
(Abels (2003)) and the distinctness condition on linearization in phonology (Richards
(2010)).
of language is composed in a parallel fashion, we predict that there should be a mechanism that introduces features at the LF side.\(^6\)

Therefore, we assume that the feature-suppletive operation is applied to event-semantics representations at the context-inference interface. To clarify this assumption, let us review Reinhart’s (2006) conception of research on the interface. She points out that for work on the interface it is useful to decompose the components of the C-I interface into three sets of systems, as depicted in (15).

\[(15)\]

The concepts systems involve the theta system as their central system. The theta system enables basic information in the concepts systems to be coded in a form legible to other systems (e.g. θ-feature clusters) on lexical items, which are the building blocks of the CS derivations. The inference system is essentially logic, and its inventory includes logical relations, functions, abstract predicates, variables, and so on. The representations generated by the CS are transferred to the inference system, which can read them as propositions suitable for its computations. The context systems narrow the information transmitted through the derivations of the CS, and select the information that is useful for the context of use. Reinhart (2006) goes on to suggest that the context and inference systems may have direct interfaces, rather than each negotiating only with the CS. We pursue this suggestion and assume the following model of interfaces.

\[(16)\]

Furthermore, Reinhart (2002) assumes that when the syntactic structure is mapped onto the inference system, it is converted into the event-semantics representation of θ-clusters. Based on these assumptions, we suppose that the feature-suppletive operation is applied to the event-semantics representa-

---

\(^6\) We are grateful to Nobuhiro Miyoshi (personal communication) for bringing this idea to our attention. However, an anonymous reviewer points out that this claim must be examined more carefully. We leave this problem for future research.
Next, we assume that the feature-suppletive operation is a kind of repair or last-resort mechanism activated only when the representation fails to meet the need of the inference system, that is, the need of ICD (Reinhart 2006). Specifically, the operation is motivated only when the event-semantics representation includes two indistinct θ-clusters which are not legible to the inference system.

Finally, we surmise that the feature-suppletive operation does the bare minimum of modification to adjust the representation to the needs of ICD. As one way to implement this view, we assume that the operation can add only the +feature. Specifically, when we find information that there exists a certain property in a context, our context system can mark it with the +feature. However, even if we perceive information that a property does not exist in a context, our context system cannot mark it with the −feature. This is based on the idea that we cannot be sure, at all times, that a certain property is not present, because even if we perceive it to be absent in a context, in the next moment we may find it to be present in that context updated with new information. Furthermore, if our context system were to mark the absence of a certain property with the −feature, the system would have to assume an additional process of rewriting this −specification when we find that the property is present in an extended context. Such an additional process will be incompatible with the idea of the bare minimum modification. Therefore, even if we find that a certain property does not exist in a context, our context system does not mark it with the −feature but leaves it unspecified to be able to mark the presence of the property with the +feature whenever we find it in an updated context.8

3.2. An Analysis with the Repair Strategy on the C-I Interface

Along the line of reasoning above, let us first discuss sentences in which the [+c] argument can co-occur with the [−m] argument. Consider the following derivation:

---

7 One might suspect that the insertion of features at LF violates the Inclusiveness condition proposed by Chomsky (1995). However, we assume that the insertion of features takes place at the context-inference interface, in an area where the syntactic computation has been completed. Thus, the problem will not arise.

8 We are most grateful to Professor Yoshiaki Kaneko and Professor Etsuro Shima (personal communication) for bringing this idea to our attention. However, we leave for future research the question of what principle underlies the assumption.
(17) \( ?[\text{TP} \text{Albert}_{+c} \text{ worried John}_{-c+m} \text{ about the adequacy of his insurance coverage}_{-m}] \).

a. \( \exists e (\ldots \& [+c](e, \text{Albert})\ldots \& [-m](e, \text{the adequacy of his insurance coverage})) \)

b. \( \exists e (\ldots \& [+c+m](e, \text{Albert})\ldots \& [-m](e, \text{the adequacy of his insurance coverage})) \)

The representation in (17a) contains the [+c] and [−m] clusters. These clusters are indistinct under our definition of distinctness, because there is no feature for which they are counter-specified. These clusters, as they stand, are illegible to the inference system. Thus, the repair strategy applies at the context-inference interface. As illustrated in (17b), if the [+c] cluster is supplied with the +m-feature from context, the cluster becomes a [+c+m] cluster, which is interpreted as an agent. As a result, the [+c+m] and [−m] clusters become distinct and legible to the inference system.

The proposed analysis can also explain why the co-occurrence of the cause role with the subject-matter role is impossible. Let us consider the following derivation:

(18) *\([\text{TP} \text{The doctor’s letter}_{+c} \text{ worried Lucie}_{-c+m} \text{ about her health}_{-m}]\).

a. \( \exists e (\ldots \& [+c](e, \text{the doctor’s letter}) \& [−c+m](e, \text{Lucie}) \& [-m](e, \text{her health})) \)

b. \( \exists e (\ldots \& [+c+m](e, \text{the doctor’s letter}) \& [−c+m](e, \text{Lucie}) \& [-m](e, \text{her health})) \)

c. \( \exists e (\ldots \& [+c](e, \text{the doctor’s letter}) \& [−c+m](e, \text{Lucie}) \& [+c−m](e, \text{her health})) \)

The representation in (18a) is insufficient for the needs of the inference system, because it includes the [+c] and [−m] clusters which are not legible to the system. To make up for this imperfection, the feature-suppletive operation is applied. There are two ways to adjust it to match interface conditions. The first is to insert the +m-feature into the [+c] cluster, as shown in (18b). However, insertion does not take place because the doctor’s letter is inanimate and hence is incompatible with the +m-feature. As a result, the inference system would not interpret (18b) appropriately. Next, let us consider the option in (18c), where the context system adds the +c-feature to the [−m] cluster. In this case, our definition of distinctness predicts that the resulting [+c−m] cluster remains indistinct from the [+c] cluster, because the shared c-feature bears the same value. Thus, this pair also fails to meet the needs of the inference system.

So far, we have shown that our approach captures in a principled manner the fact that the human subject of the verb worry can co-occur with
the subject-matter argument, whereas the inanimate subject cannot. A key to solving this puzzle is the intentionality of the subject. The presence or absence of intentionality is also observed among animate arguments. If so, our analysis makes further predictions: if the animate subject is interpreted as just a cause without intentionality, its co-occurrence with the [−m] argument is not allowed. This prediction is borne out by the following data obtained from our informant.

(19) a. The math teacher bothered Lucie about a math problem.
   (Under agentive reading: natural) (Under causative reading: unnatural)
   b. The math teacher bothered Lucie about a math problem intentionally.

We disambiguate the two interpretations in (19) by adding certain adverbs. In (19b), the adverb intentionally forces the agentive reading. In contrast, introducing the adverb accidentally, we obtain the unnatural causative reading in (19c).

Finally, let us consider the contrast between the lexical and analytic causatives that Reinhart (2002) argues is motivation for the application of CD at the lexicon-syntax interface. Since ICD is imposed after syntactic computations, it operates on a larger unit than the word. Therefore, we assume that ICD is applied to a full TP on the basis of the assumption that a full TP is a cyclic node (Reinhart and Siloni (2005)). This assumption means that the projection of a full TP (i.e. the projection of [Spec, TP]) is necessary to complete a cycle, where every θ-role of a verb must be discharged, so every time the CS generates a full TP, it may be sent to LF (cf. Chomsky (2004), Quicoli (2008)). Given these assumptions, ICD is applied cyclically at each transferred TP, as illustrated in (20).

(20) [TP The doctor’s letter [+c] made [TP Lucie[−c+m] worry about her health[−m]]].

Furthermore, we assume with Ritter and Rosen (1993) that an analytic causative selects a complement headed by an inflectional category, that is, a TP complement. With these in mind, let us next look at the analytic causative in (21).

(21) [TP The doctor’s letter [+c] made [TP Lucie[−c+m] worry about her health[−m]]].

Under this cyclic Spell-Out model, ICD is applied to the embedded TP, and
then to the matrix TP. A crucial point is that the [+c] and [−m] clusters are not evaluated in the same TP domain in contrast to the lexical causative and thus they do not form a pair illegible to the inference system. Therefore, the proposed model accounts for the analytic causative. ⁹

4. Conclusion

We have argued that Reinhart’s (2002) theta system fails to capture the fact that the agent role can co-occur with the subject-matter role while the cause role cannot, and have tried to account for the co-occurrence facts with a legibility condition of the inference system and a repair mechanism applied at the context-inference interface. Furthermore, our approach, in tandem with a cyclic Spell-Out model, has given an explicit explanation to the contrast between the lexical and analytic causatives.

REFERENCES


⁹ We leave an analysis of the analytic causative have for future research.


[received April 1 2012, revised and accepted July 23 2012]

Department of English Linguistics
Tohoku University
Kawauchi, Aoba-ku, Sendai-shi, Miyagi 980–8576
e-mail: kenji-s@sal.tohoku.ac.jp