THE "AVOID EMBEDDING" CONDITION
AND THE CHECKER/CHECKEE ASYMMETRY

HIROYUKI TANAKA
Osaka University*

1. The "Avoid Embedding" Condition Revised

In order to derive the effect that strong features cause strict cyclicity, Chomsky (1995: ch. 4) proposes that strong features have to be eliminated before the projection containing them is embedded by a projection of another category. This can be stated as follows:

(1) D is canceled if α is in a category not headed by α,
[where D=derivation, α=category with a strong feature].
(Chomsky (1995: 234))

From (1) it follows that any element that can be embedded by a projection of another category must not have a strong feature. The structure in (2) illustrates this. Here, X with a strong feature F gets embedded by another category ZP before F is checked off, and the derivation is canceled at this point.

(2)

```
   *ZP
  /   \
Z X   P
  |   |
X F  YP
 [F]strong
 ...WP...
 [F]
```

* This is part of the paper presented at a workshop in the 22nd annual meeting of Kansai Linguistic Society, held at Kyoto University on November 8th, 1997. I would like to thank the participants and two anonymous EL reviewers for their useful comments.

1 This is in turn deduced from:

(i) Nothing can join to a nonprojecting category. (Chomsky (1995: 234))
In this squib I claim that empirically (1) is both too strict and too weak, and propose that (3) is the correct description.

(3) D is canceled if $a$ is in a category not headed by $a$, where $a=$category with a checker.

Here I use the technical term *checker* to refer to the feature on the target in a checking relation (following the terminology of Chomsky (1995: 278)), as opposed to the corresponding *checkee* attracted into its checking domain (i.e. specifiers (4a) or the head-adjoined positions (4b)).

(4) a. $\text{checkee} \Rightarrow WP$

\hspace{1cm} b. $\text{checker} \Rightarrow X^0$

\hspace{1cm} $\text{checker} \Rightarrow X$

\hspace{1cm} $\text{checkee} \Rightarrow W$

\hspace{1cm} $\text{checkee} \Rightarrow X^0$

According to (3) the derivation is canceled if it builds up a structure like (5), where a checker is embedded before it is checked off. (3) is both more and less restrictive than (1), and each aspect brings about different empirical consequences.

(5) $^*ZP$

\hspace{1cm} Z

\hspace{1cm} XP

\hspace{1cm} X

\hspace{1cm} YP

\hspace{1cm} [F]_\text{checker}

\hspace{1cm} [F]

Before we go on to examine the consequences, we have to validate the use of the notions *checker* and *checkee*. Conceptually, the distinction is not necessary because there is no a priori reason to claim that asymmetry exists between the features involved in a checking relation. But empirically, it seems that which role an element plays in the asymmetric relationship is determined by UG. Just as it is universally determined that DP has a Case feature, it is also determined that the Case feature of DP is a checkee (or assignee, in traditional terms), not a checker (assigner). Thus, we want to prevent the obviously un-

2 Read “checker $\Rightarrow X$” in (4a, b) as “checker belongs to X.” I assume with Chomsky (1995: ch. 4) and Collins (1997) that all checkers are $-\text{Interpretable}$.

3 For convenience I use the expression “a feature F is embedded” to mean “the category containing F is embedded by a projection of a different category.”
wanted result that DP checks Case with another DP and V or T checks Case with another V or T. As Chomsky (1995: 278) notes, the asymmetry between checkers and checkees "play[s] a certain role in computation." Having no principled account at hand of how we can determine whether a given feature is a checker or a checkee, suffice it here to say that UG prescribes that DP have a checkee Case feature and V and T a checker Case feature, and likewise for other types of features.

2. The Mechanism

Notice that (1) and (3) are the same in that they both constrain the embedding of features, but they differ as to what kind of feature does not tolerate embedding. (3) refers only to checkers, leaving checkees exempt from the "avoid embedding" condition. This "weakening" aspect of the revision allows for the possibility that a strong feature can be embedded if it is a checkee. Consider argument DPs for example. They must be embedded in a structure (under a projection of predicates, perhaps) to be thematically interpretable at all. The revised system allows a DP argument with a strong feature to be embedded until a functional head with an appropriate feature gets merged with the structure. This is a desirable consequence because, as we will show in the next section, the widely observed phenomenon of "weak pronoun shift" strongly suggests that weak pronouns are possessed with a strong feature.

The other, "strengthening" aspect of (3) is that it requires that all —interpretable features, as long as they are checkers, be eliminated before embedded. This leads us to the interesting prediction that no head can check its checker features in its raised position. It should be noted that this consequence is irrelevant to Chomsky's original aim in proposing (1), but it is nevertheless a fact that (1) does not itself deduce this effect, nor does any other mechanism in Chomsky (1995: ch. 4). Another theoretical implication of the revised condition is that it forces on us the overall architecture of syntactic computation in which

---

4 I use the symbol V as a verbal element which is responsible for accusative Case checking, abstracting away from the v-V distinction in Chomsky (1995: ch. 4).

5 Watanabe (1997) independently reaches the conclusion that strong features can move in overt syntax.
there is only one cycle in the derivation, contra Bures (1993) and Watanabe (1995). In order for the derivation not to be canceled, even weak features have to attract some feature before they are embedded. However, we do not want them to cause overt movement. We are therefore led to the view that attraction involving only weak features does not pied-pipe the whole category, while that involving a strong feature does (see Bobaljik (1995) and Groat and O’Neil (1996)).

3. Embedding Strong Checkees Allowed

Let us see the evidence that makes Chomsky’s formulation in (1) too strong. As an instance of this we will consider what may be called “weak pronoun shift” and its theoretical implications. Generally across languages, weak (i.e. unstressed) pronouns tend to be placed higher in the syntactic structure than do other noun phrases. One of the clearest illustrations is pronominal clitics.6 Take for example the cliticized pronoun in French functioning as the direct object of the verb.

(6)

a. Jean aime Marie et Charlotte.
   ‘Jean loves Marie and Charlotte.’

b. Jean les aime.
   ‘Jean loves them.’

Assuming with Kayne (1975) that this kind of displacement is caused by syntactic movement,7 the problem relevant here is what triggers the movement. If in our framework strong features are the cause of overt movement, we want to say that the clitic has a strong affixal feature that requires it to be attached to an appropriate head.

Faced with this situation, Chomsky’s (1) is too strong in prohibiting the embedding of strong features under a projection of different categories: If a clitic bearing a strong feature is introduced in the structure, it must not be embedded before the feature is checked off, but at the same time it must be embedded by a projection of a predicate before

---

6 Other examples include object shift affecting only weak pronouns in languages such as Swedish, the English Verb-Particle constructions taking pronominal objects, and the impossibility of weak pronouns to induce singular concord in Belfast English (Henry (1995)).

7 See Borer (1986) for an overview of movement vs. non-movement approaches to clitic pronouns.
the checking takes place in order to be interpreted as an argument. One might propose that this conflict be resolved by assuming that the host (target) of cliticization has the strong feature to attract the clitic, thereby allowing the clitic to have a weak feature. But this move would obscure the insight that the cause of movement is on the part of the pronoun, and would also fail to capture the locality of clitic movement, as shown in the impossibility of moving an object clitic to a higher, non-restructuring verb in Italian (see Rizzi (1982: ch. 1)).

Our (3), on the other hand, readily allows the clitic to have a strong feature and to be embedded in the structure. (3) states that every checker must not be embedded. Since the pronominal clitic is an argument DP, its affixal feature is a checkee and therefore it can be embedded. Since it is strong, the movement is overt.

4. Embedding Weak Checkers Disallowed

There are also some pieces of empirical evidence that make Chomsky’s formulation in (1) too weak. Our proposal serves to capture a systematic gap in natural language, which is too prevailing to be even noticed so far, as an automatic consequence of (3). (3) entails that every checker has to accomplish checking before it is embedded, and hence within its own projection. In other words, it is impossible for a checker feature to wait until it is embedded by another category and to undergo checking after the head containing it has raised to a higher head. Thus we predict the generalization in (7), which as far as I

8 An anonymous EL reviewer pointed out that the motivation of the clitic movement in (6b) may be attributed to Kayne’s (1994) Linear Correspondence Axiom (LCA), as suggested in Chomsky (1995: 337). It is, however, highly dubious that this is the sole motivation of clitic movement. The relevance of the LCA to clitics is that a clitic, being a non-complex DP under the bare phrase structure theory, violates the LCA if it is generated in a complement position and remains there. If it is generated elsewhere, however, the LCA becomes irrelevant and cannot force the movement. (i) represents a case in which the clitic is generated as the subject of a small clause, which cannot be considered as a complement; but the clitic still has to move.

(i) a. Jean lesi croit [ti contents de leur chambre].
   ‘Jean believes them satisfied with their room.’
   b. *Jean croit [les contents de leur chambre].

9 This analysis lends itself to the claim that the operation Move exists, in the sense that the operation is triggered by the moving element and not by the target.
know is exceptionless crosslinguistically.

(7) No head raises to have its checker features checked off. In general, elements with checkee features (DPs, wh-phrases, etc.) usually move up to the checking domain of the target, and this is allowed in our system, as seen in section 3. What is banned is the movement of the head to check its checker features against the checkees in the checking domain of the upper head it has adjoined to.

In terms of Case, (7) states that a DP may raise to get Case from higher head, but no head may raise to assign Case to a higher DP. It is safe to say that this is a correct generalization. The assignment of Case to DP may involve movement of DP into the domain of the Case-assigner (either overtly (8) or covertly (9)), but no cases of Case-assignment have been attested which crucially involve movement of the Case-assigner to the vicinity of the assignee, as evidenced by the lack of sentences like (10).

(8) He is believed [t, to be a genius].
(9) There is likely to be a riot around here.
(10) *It is likely [TP himi to [vP ti mention that the earth is flat]].
    (In the sense ‘He is likely to mention that the earth is flat.’)

Suppose that in (10) him, the logical subject of mention, has moved up to the specifier of the embedded TP to satisfy the EPP requirement. The verb mention can have an accusative Case feature even when it takes a that-clause complement10 (note its ability to Case-mark the expletive it as in They never mentioned it to the candidate that the job was poorly paid; see the discussion in Authier (1991)). Without the restriction in (7), which is derived from (3), it is possible that the verb moves up in covert syntax to the infinitival T to check its Case feature (a checker) against that of him (a checkee), leading the derivation to convergence. This can be avoided if the Case feature of mention does not tolerate embedding, as required by (7).

An EL reviewer points out that it would have to be inserted into the Spec, TP instead of moving him (thus generating (11), which wins over

10 To be more precise, I assume that it is the light verb (v) of the split VP structure that is responsible for accusative Case checking (Collins (1997), Ura (1996)), and that Case checking does not take place in θ-positions (Chomsky (1995: ch. 4)). These assumptions reject the possibility (pointed out by the reviewers) that him is Case-checked in its base position (and hence inside vP), which would make (7) irrelevant in ruling out (10).
(10)) in Chomsky's system of economy where Merge is preferred to Attract/Move.

(11) *It1 is likely [TP t1 to [vP him mention that the earth is flat]].

Then we want this derivation to crash in order for (10) to be relevant to our discussion. For the lack of space I cannot present a detailed discussion of this case. A solution to this problem I have in mind is that the "more economical" derivation does not converge precisely because (7) prohibits Case checking between him and mention, and that this is the only way to rule (11) out under the present assumptions.

This system disallows Case-assigning heads to be stacked up together before their Case features are eliminated. It therefore rejects Chomsky's mechanism of LF accusative Case feature checking, in which the formal features of the verb and the object enter into checking relations in the checking domain of T.\textsuperscript{11} Without the restriction in (7), it is also unclear how to block the unwanted derivation in which the Case features of nominative-assigning T and accusative-assigning V, for instance, are wrongly associated with the logical object and subject, respectively, when V has overtly adjoined to T (e.g. in French) before finishing accusative Case checking.

It is also important to notice that (7) has a direct implication for the phenomena of Case absorption. If all Case checkers have to undergo checking before embedded, and if Case absorption is determined by the external context of the projection of the Case assigning head (Watanabe (1996)), one needs a sophisticated theory of the interaction between selection and the status of Case features. In other words, we must at least leave open the possibility that the property of checkers (i.e., whether they are absorbed or not) is affected by what kind of projection they are immediately embedded under.

Let us note finally that the restriction in (7) has another important theoretical implication, namely that it imposes an Agr-less system of Case checking. Since a Case assigner has to check its Case within its own projection, the Agr-mediated system of Case checking (Chomsky (1993)), where V and T have to raise to Agr to check Case features,

\textsuperscript{11} Thanks to an EL reviewer for reminding me of this point. He/she also mentions that my proposal runs counter to Miyagawa's (1997) analysis of Japanese clause-internal scrambling, in which objects enter into Case checking with V within the checking domain of T.
can no longer be maintained. Whereas Chomsky (1995: ch. 4) rejects the existence of Agr as an independent syntactic head on a purely conceptual ground, our argument here suggests another way to do so, though in a rather technical fashion.

REFERENCES


4-2-6-101 Minami-Hanayashiki
Kawanishi-shi
Hyogo 666-0026
e-mail: htanaka@osk.3web.ne.jp