Kase as a Weak Head[*]

Mamoru Saito
Nanzan University

SUMMARY

Chomsky (2013) proposes that a finite clause can be labeled because of φ-feature agreement. This raises a question on languages like Japanese that lack φ-feature agreement. I suggest in this paper that the KP (Case phrase) hypothesis, argued for in Travis and Lamontagne (1992), opens up a way to answer this question. More specifically, I propose that Ks (Case markers) in Japanese are weak heads in the sense of Chomsky (2015), and hence, cannot participate in labeling. The proposal predicts that finite clauses of the form, {{DP, K}, TP}, are labeled by T. Another consequence is that although Case-marked DPs have the form, {DP, K}, D provides their label.

RÉSUMÉ

Chomsky (2013) propose qu'une proposition finie puisse être étiquetée en raison des accords de trait f. On se pose donc des questions sur des langues qui en manquent, comme le japonais. Je suggère dans cet article que l'hypothèse SK (syntagme de cas), avancée par Travis et Lamontagne (1992), permet de répondre à cette question. Plus précisément, je propose que les K (marqueurs de cas) en japonais soient des têtes faibles dans le sens de Chomsky (2015); ils ne peuvent donc pas participer à l'étiquetage. Cette suggestion prédit que les propositions finies de la forme {{SD, K}, ST} sont étiquetées par T. Une autre conséquence est que, bien que les SD marqués par le Cas aient cette forme, {SD, K}, D fournit l'étiquetage.

1 INTRODUCTION

Travis and Lamontagne (1992) was one of the first works to argue for the Case phrase (KP) hypothesis. According to this, the object DPs in the Japanese (1) and the English (2) have the structures in (3a) and (3b) respectively.

---

(1)   Hanako-ga   Taroo-o   sikat-ta.
        Hanako-NOM Taroo-ACC scold-Past
        'Hanako scolded Taroo.'

(2)   Mary scolded John.

(3)   a.   KP   b.   KP
               /   \   /   \
            DP   K   K   DP
             | | | |
           Taroo   o   Ø   John
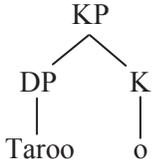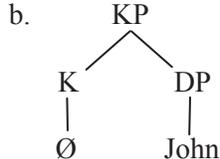
    In this paper, I argue that this KP hypothesis leads to a solution for a problem that arises in my recent work on labeling in Japanese. Chomsky (2013) proposes that a finite clause with DP in the subject position, {DP, TP}, is labeled through ϕ-feature sharing between D and T. On the assumption that Japanese lacks ϕ-feature agreement, I suggested in Saito (2016), for example, that suffixal Cases in the language function as anti-labeling devices and make phrases invisible in the labeling process. Then, βP provides the label for γ in (4).

(4) $\gamma = \{\alpha\text{P-Case}, \beta\text{P}\}$

As the subject DP appears with Case, as shown in (1), it is predicted that TP provides the label for a finite clause in Japanese, which is of the form {DP-Case, TP}. This analysis, if correct, raises a new question: why is it that Case functions as an anti-labeling device? Adapting Chomsky's (2015) distinction between strong and weak heads, I entertain the possibility in this paper that this is because suffixal Cases in Japanese are weak heads that are unable to participate in labeling.

    In the following section, I first review Chomsky's (2013) labeling algorithm. Then, I introduce the proposal on Japanese in Saito (2016) and explain the problem it raises. In Section 3, I illustrate how the extension of the KP hypothesis solves the problem. This involves a slight reinterpretation of Chomsky's (2015) weak heads. Section 4 concludes the paper.

2   LABELING IN JAPANESE

2.1   CHOMSKY (2013) ON LABELING AND ϕ-FEATURE AGREEMENT

Among the fundamental questions that have been pursued in the minimalist approach is why ϕ-feature agreement is observed in languages. In this section, I briefly go over Chomsky's (2013) proposal on this. I first describe the mechanism of ϕ-feature valuation he assumes, and then present his labeling algorithm.

   Chomsky (2008) proposes that the valuation of ϕ-features on T, for example, takes place as in (5).

(5)

```
        TP                    →              TP
       /  \                                 /  \
      T    vP                              T    vP
   [φ: _ ]   \                          [φ: α]   \
            DP    vP                            DP    vP
          [φ: α]                              [φ: α]
          [Case: _ ]                          [Case: NOM]
```
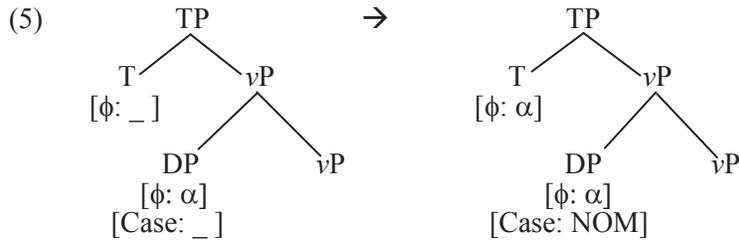
T, with unvalued φ-features, searches its domain and enters into Agree relation with the subject DP. T obtains the φ-feature values from the DP and as a reflex of this, the Case feature of the DP is valued as nominative.

Then, why is it that φ-feature agreement exists in the first place? Chomsky (2013) provides an answer with his labeling algorithm. He notes first that syntax not only forms a constituent with Merge as in $\gamma = \{\alpha, \beta\}$, but must also specify the label (nature) of γ. The idea is that when a verbal element and a nominal element are merged, the interpretation of the constituent depends on whether it is verbal (VP) or nominal (NP). Hence, a constituent must be labeled when it is transferred to the interfaces. Given this, he considers the following three cases of Merge:

(6) a. $\gamma = \{H, \alpha P\}$
     b. $\gamma = \{\alpha P, \beta P\}$
     c. $\gamma = \{H_1, H_2\}$

The case (6a) is straightforward as search into γ yields a unique head H. Then, it can be assumed that H provides the label of γ in this case. On the other hand, the label cannot be determined in this way for (6b) and (6c).

However, Chomsky notes that (6b) arises in actual derivations and examines what happens in those cases. Let us take a simple transitive sentence as in (2). Its derivation is illustrated in (7).

(7)

```
        YP  <φ, φ>
       /  \
      DP    TP
      [φ]  /  \
          T    XP (vP)
         [φ]  /  \
             DP    vP
             [φ]  /  \
                 v    VP
                    /  \
                   V    DP
```
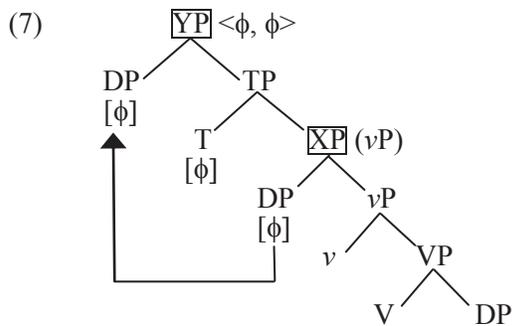
The structure is constructed bottom-up and the configuration in (6b) first arises when the subject DP merges with vP. Here, the DP moves out of XP after T is introduced into the structure. Chomsky assumes then that vP provides the label for XP because it is the only element that is fully contained within XP. But the subject DP merges with TP, creating the configuration in (6b) again. This is where φ-feature agreement comes into play. The search into YP yields DP and TP. Further search into these two phrases locates two heads D and T with identical φ-features. Chomsky proposes that YP can be labeled as <φ, φ> because of this feature sharing.

The analysis illustrated above not only constitutes a hypothesis on why languages have ϕ-feature agreement but also accounts for the distribution of DPs to a large extent, including the last resort nature of NP-movement and why it applies obligatorily when it does. In (7), for example, the subject DP must move out of XP for otherwise XP fails to be labeled. And NP-movement can terminate only in a position of ϕ-feature sharing.

2.2 SUFFIXAL CASE AS AN ANTI-LABELING DEVICE

The theory outlined above raises a number of questions for the analysis of languages like Japanese, which lacks ϕ-feature agreement at least on the surface. For example, according to Chomsky, Case is valued as a reflex of ϕ-feature agreement. How then is Case valued in Japanese? Further, a finite clause can be labeled as <ϕ, ϕ> because of ϕ-feature agreement. How are finite clauses in Japanese labeled then? These are the questions I addressed in Saito (2016).

I first adapted Bošković's (2007) proposal that Case valuation takes place independently of ϕ-feature valuation. His analysis is illustrated in (8).

(8)

<div style="display:flex">
TP → TP → TP (<ϕ, ϕ>)
</div>

Tree 1: TP branches into T [ϕ: _] and vP; vP branches into DP [ϕ: α] [Case: _] and vP.

Tree 2: TP branches into T [ϕ: α] and vP; vP branches into DP [ϕ: α] [Case: _] and vP.

Tree 3: TP (<ϕ, ϕ>) branches into DP [ϕ: α] [Case: NOM] and TP; TP branches into T [ϕ: α] and vP; vP branches into __ and vP.
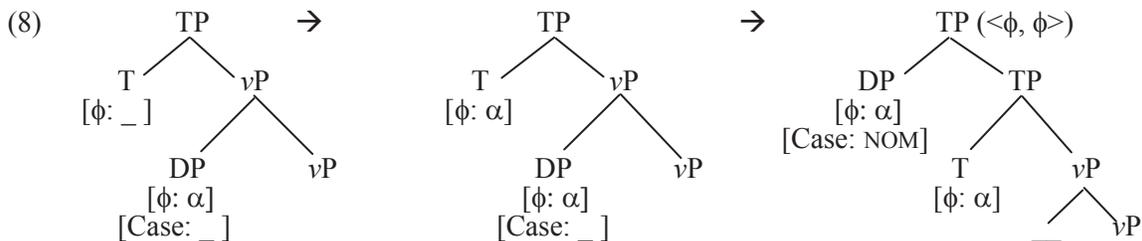
The ϕ-features of T are valued exactly as Chomsky (2008) proposes, but the Case feature of the subject is not valued at this point, as the middle structure illustrates. According to Bošković (2007), the subject DP raises to Spec,TP, as shown in the rightmost structure, and probes T so that its Case feature can be valued as nominative.

Noam Chomsky and Hisa Kitahara point out that the Case feature is not valued through a probe-goal relation because DP is not a head. I assume, following Kitahara's suggestion, that the Case valuation is achieved through search into {DP, TP} as in (9).

(9)

TP (<ϕ, ϕ>) branches into DP and TP. DP branches to D [Case: _] [ϕ: α]. TP branches to T [ϕ: α].
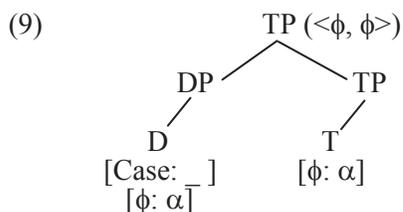
Search into the structure finds DP and TP. Further search into both yields D with an unvalued Case feature and T. This makes it possible for T to value the Case feature on D. If this is tenable, Case feature valuation is accomplished by exactly the same mechanism as the labeling of the clause as <ϕ, ϕ>. With or without this revision, Bošković's (2007) proposal on Case feature

valuation is applicable to Japanese as it does not rely on ϕ-feature agreement. The nominative Case on the subject *Hanako* in (1), for example, can be valued exactly as in (9).

The next question to be addressed is how finite clauses are labeled in Japanese. Recall that (9) is labeled as $<\phi, \phi>$ because D and T share identical ϕ-features. But if Japanese lacks ϕ-feature agreement, then there are no ϕ-features on T in the language. Then, a finite clause {DP, TP} fails to be labeled. Here, I proposed in Saito (2016) that suffixal Cases in Japanese function as anti-labeling devices. In other words, Case markers make DPs invisible in the labeling process. This allows XP to provide the label for γ in (10).

(10) $\gamma = $ {DP-Case, XP}

This is consistent with the fact that Case-marked elements never "project" further, and allows Japanese finite clauses to be labeled. The Japanese counterpart of (9) is not labeled $<\phi, \phi>$ but is labeled by the lower TP because the subject DP accompanies a Case marker.

Further, the proposal accounts for why scrambling is possible in Japanese, as in (11).

(11)   Taroo-o   Hanako-ga   __   sikat-ta.
         Taroo-ACC Hanako-NOM   scold-Past
         'Hanako scolded Taroo.'

In this example, the object DP, *Taroo-o*, internally merges with TP as in (12).

(12)   ??

         DP-ACC   TP

As this is an {XP, YP} structure without feature-sharing, scrambling is illicit, for example, in English. But thanks to the suffixal Case on the object DP, the structure is allowed in Japanese. As the accusative Case marker makes the object DP invisible for labeling, TP provides the label for the structure.

This analysis has other desirable consequences as well, as discussed in Saito (2016), but raises a new question. Why is it that suffixal Case markers in Japanese function as anti-labeling devises? In the following section, I suggest that the KP hypothesis leads to an answer.

3   KASE IN LABELING

3.1   ON THE STATUS OF CASE MARKERS IN PHRASE STRUCTURE

I first briefly introduce Chomsky's (2015) distinction between strong and weak heads. Then, I consider the labeling issue with the KP hypothesis on the assumption that K is a weak head.

Chomsky (2015) addresses Comp-trace effects and the pro-drop parameter. For the latter, he proposes that the parametric variation exists because pro-drop languages allow the structure in (13a) whereas non-pro-drop languages do not.

(13)   a.   TP   b.   TP ($<\phi, \phi>$)

                 T vP   DP   TP

                                                                  T vP
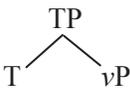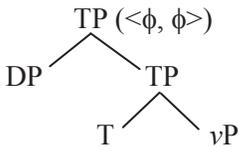
The hypothesis is that T, being weak, does not qualify to provide the label for (13a) in non-pro-drop languages. Then, the only way that a finite clause can be labeled is by merging the subject DP with TP as in (13b). (13b) is labeled as $<\phi, \phi>$ because of ϕ-feature sharing. Then, a finite clause can be labeled only when the subject is present in Spec,TP. This, according to Chomsky, is the EPP. It is assumed here that ϕ-feature sharing makes T strong, and consequently, T labels the lower TP in (13b).

    In this section, I assume, following Travis and Lamontagne (1992), that Case markers in Japanese are K heads, and suggest that they are weak. In addition, I slightly reinterpret the labeling mechanism involving weak heads. Chomsky assumes that {H, XP} simply fails to be labeled when H is weak. I suggest instead that search for a label provider applies to XP in this case. The proposal is stated more precisely in (14).[1]

(14)   Search {α, β} for a label. If α is a weak head or search into α yields a weak head, then search on the α side is suspended and it continues only on the β side.

Let me illustrate this with the examples in (15).

(15)   a.   {DP, K}
        b.   {{DP, K}, {vP, T}}

(15a) is the structure of a Case-marked argument. Search into this structure immediately yields a head K. Given that K is weak, search for a label provider shifts to DP and eventually locates D. This is a desirable result. When a Case-marked DP appears in the object position, for example, V is in selectional relation with D, not K. Hence, D should provide the label for the object.

    (15b) is the configuration of a Case-marked DP in the subject position. The structure is illustrated in (16).

(16)   ?? (TP)

           DP   TP

       DP   K vP   T
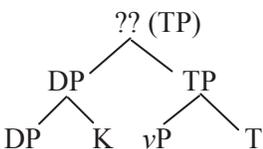
Search into {DP, TP} applies to seek a label provider. Further search into DP locates the head K. Since it is weak, the search shifts to the alternative, that is, TP. On the assumption that T is a strong head in Japanese, T provides the label for (16). Here, search into DP and TP does yield the two heads, K and T. This allows the Case feature of K to be valued as nominative. It is just that K does not count for labeling as it is a weak head.

---

[1] Given (14), (13a) is labeled by vP when T is weak. I assume that this is incompatible with the selectional requirement of C, which (13a) later merges with.

3.2   EXTENSIONS TO SCRAMBLING AND MULTIPLE SUBJECT SENTENCES

The proposal in (14) extends to labeling of sentences with scrambling and multiple nominative subjects. Let us start with the former.

The example in (11) is repeated below in (17).

(17)   Taroo-o   Hanako-ga   __   sikat-ta.
        Taroo-ACC Hanako-NOM   scold-Past
        'Hanako scolded Taroo.'

The relevant part of the structure of this sentence is as in (18).

(18)   ?? (TP)

              DP   TP (= (16))

        DP   K   DP   TP

               | o DP   K *v*P   T
               o |
                                ga
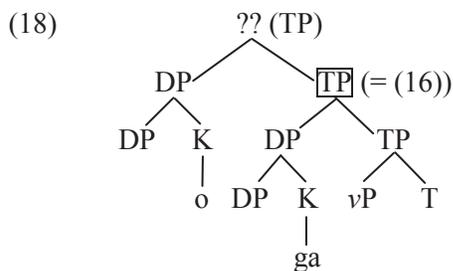
In this structure, the object DP is internally merged with a finite clause, which is labeled as TP as illustrated in (16). The issue is how the resulting structure is labeled.

Search into (18) first yields DP and the boxed TP. A weak head K is found with further search into DP. Thus, search shifts to the boxed TP. One finds the subject DP and the lower TP in this case. Further search into the subject DP yields a weak head, K, again. Search shifts to the lower TP and a label provider T is finally located. Thus, the whole structure is labeled by T. This accounts for why scrambling is allowed in Japanese.

It is suggested in Saito (2016) that Bošković's (2007) mechanism for Case valuation, together with the analysis of suffixal Cases as anti-labeling devices, leads to an explanation for why multiple subject sentences are allowed in Japanese. The most famous example from Kuno (1973) is shown in (19).

(19)   Bunmeikoku-ga   dansei-ga heikin-zyumyoo-ga   mizika-i.
        civilized.country-NOM male-NOM average-life.span-NOM short-Pres.
        'It is in civilized countries that male's average life span is short.'

In the remainder of this section, I show that the suggestion can be made precise with the KP hypothesis.

Multiple subject sentences have a structure identical to (18), except that the initial DP accompanies nominative instead of accusative Case. The structure is shown in (20).

(20)

```
              ?? (TP)
          DP          TP (= (16))
      DP   K     DP       TP
           |    DP  K   vP   T
          ga         |
                    ga
```
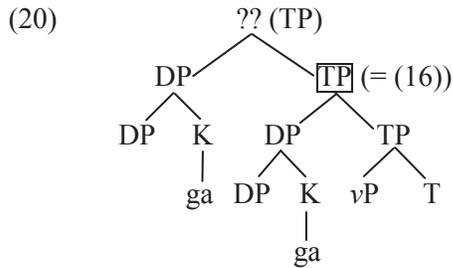
Then, labeling can be achieved exactly as in the scrambling case. Initial search yields the higher subject DP and the boxed TP. As the former is headed by K, the boxed TP is the target for further search. This brings us to the lower subject DP and the lower TP. As search into the subject yields K, search shifts to TP and locates the label provider T. Thus, (20) is successfully labeled by T.

Here, a slight adjustment is necessary for the valuation of Case of the higher subject. I argued above that nominative Case on the subject is valued in the configuration of (16), repeated below in (21).

(21)

```
            ?? (TP)
         DP        TP
      DP    K    vP    T
```
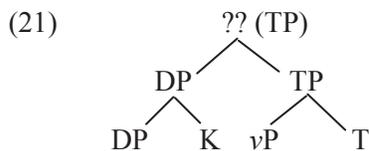
Search into the structure finds DP and TP. Then, further search into both locates K and T. This allows K to be valued by T as nominative. However, the situation is a little more complex in the case of higher, extra subjects. This is best illustrated with examples with three subjects as in (19). The structure is as follows:

(22)

```
                  TP
            DP            TP
         DP   K     DP         TP (= (16))
              |   ga DP  K   DP      TP
             ga        |   ga DP  K vP   T
                               |
                              ga
```
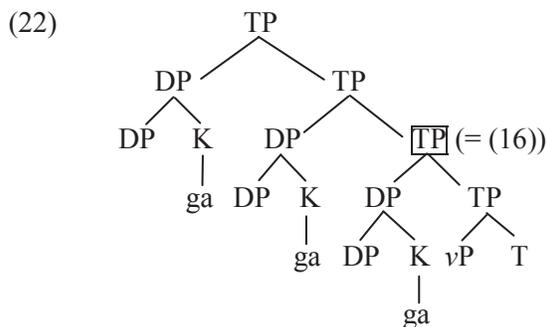
Search into this structure yields the topmost subject DP and its sister TP. Search into the DP locates K. On the other hand, the first head found by search into the TP is K in the second subject. Then, it is not obvious how K in the topmost subject DP is valued as nominative.

As a technical solution to this problem, I tentatively suggest that it is the label provider that values a Case feature. That is, in the configuration in (23), the label provider of ZP values the feature F.

(23)

```
                    XP
                   /  \
               YP      ZP
                  \
                   W
                 [F:_ ]
```
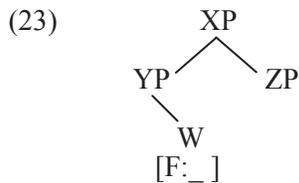
Then, K in the topmost subject DP in (22) is successfully valued as nominative by T.

4   CONCLUSION

I argued in this paper that the KP hypothesis, argued for in Travis and Lamontagne (1992), opens up a way to pursue an issue that arises with labeling in Japanese. I proposed in Saito (2016) that suffixal Case markers in Japanese function as anti-labeling devices. The issue that I discussed in this paper is why this should be so. I showed that the KP hypothesis, together with the assumption that K is a weak head, leads to a possible explanation.

Travis and Lamontagne (1992) propose that Case (K) is a head in all languages. On the other hand, I argued in this paper that if K is a weak head in Japanese, it can be explained why the language allows scrambling and multiple subject sentences. Then, if K is a weak head in all languages, scrambling and multiple subject sentences should be observed universally. Since this is not the case, it must be that in languages like English, K is either strong or absent. The former is consistent with Travis and Lamontagne (1992), but it remains to be seen whether it is tenable.

REFERENCES

Bošković, Željko (2007) "On the locality and motivation of Move and Agree: An even more minimalist theory," *Linguistic Inquiry* 38: 589-644.

Chomsky, N. (2008) "On phases," in R. Freidin, C. P. Otero and M. L. Zubizarreta, eds., *Fundamental issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133-166, Cambridge, Mass.: MIT Press.

Chomsky, N. (2013) "Problems of projection," Lingua 130: 33-49.

Chomsky, N. (2015) "Problems of projection: Extensions," in E. Di Domenico, C. Hamannand and S. Matteini, eds., *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 3-16, Amsterdam: John Benjamins.

Kuno, S. (1973) *The structure of the Japanese language*, Cambridge, Mass.: MIT Press.

Saito, Mamoru (2016) "(A) Case for labeling: Labeling in languages without φ-feature agreement," *The Linguistic Review* 33: 129-175.

Travis, L. and G. Lamontagne (1992) "The Case filter and the licensing of empty K," *Canadian Journal of Linguistics* 37: 157-174.