GOVERNMENT OF PRO, BY PRO, FOR PRO*

KATSUNORI FUKUYASU

Tottori University

In this paper, I argue that the theorem that PRO is ungoverned is generalized to the zero-level category N. I propose a condition on government and try to establish an ideal notion of government for PRO, and argue that this notion of government accounts for the marginal but significant phenomenon found in English as well as in Japanese.

1. INTRODUCTION. In the government-binding (GB) framework, Chomsky (1981, 1982, 1986a) tries to account for the distribution of empty categories which include trace and PRO. Chomsky argues that the property that PRO is ungoverned follows from the binding theory:

(1) Principles of the Theory of Binding
   A. An anaphor is bound in its governing category.
   B. A pronominal is free in its governing category.
   C. An R-expression is free.

Chomsky (1982: 20) defines the governing category for α to be the minimal S or NP containing α and a governor of α. He explains the basic property of PRO as follows: “Since PRO is a pronominal anaphor, it is subject to Principles A and B of the binding theory, from which it follows that PRO lacks a governing category and is therefore ungoverned” (Chomsky 1982: 21). In order for PRO not to have a governing category, PRO must not have a governor by the definition of the governing cate-

* The main idea of this paper originates from my “Gaps in NPs” presented at the 4th meeting of The Tsukuba English Linguistic Society at the University of Tsukuba in November 13, 1983, and a revised and extended version was read at the 93rd general meeting of the Linguistic Society of Japan held at the Osaka College of Foreign Studies, October 12, 1986. I am sincerely grateful to the participants of those meetings, in particular, Minoru Yasui, Minoru Nakau, Shosuke Haraguchi, Yukio Hirose, Masaki Sano and Taisuke Nishigauchi, for several valuable suggestions and comments. I am indebted to two anonymous English Linguistics reviewers for useful comments and suggestions. I would also like to express my thanks to Dr. P. J. Lewitt of Tottori University for his assistance.

—186—
gory. Since PRO must not have a governor, we get the following theorem:

(2) PRO is ungoverned.

Assuming that the binding theory (1) has something to do with a referential function in some sense, and that only phrasal categories have certain referential functions, PRO in 2 is understood to be a phrasal category (i.e., NP).

2. Extraterritoriality Condition. Let us begin with the notion of government assumed here. Consider first the following:

(3) a. Internal Government

\[ \beta \ldots \gamma \ldots \alpha \ldots \gamma \], where

(i) \( \alpha = X^0 \) (in the sense of X-bar theory)
(ii) where \( \phi \) is a maximal projection, \( \phi \) dominates \( \alpha \) if and only if \( \phi \) dominates \( \gamma \).

b. External Government

If \( \alpha \) internally governs \( \gamma \), \( \alpha \) still governs the head of \( \gamma \), externally.

What the definition 3a means is roughly that if \( \alpha \) and \( \gamma \) are contained in the same maximal projection, the head of the maximal projection, \( \alpha \), governs \( \gamma \). From 3b it can be inferred that if \( \gamma \) is governed, the head of \( \gamma \) is also externally governed.

The intuitive idea which I want to develop is that the head of one government cannot govern the head of another government or interfere with another government, if there are both an appropriate governor and its governees in the latter government. In 3, the head of \( \gamma \) is governed by \( \alpha \) externally, if there is no "internal government relation" within \( \gamma \). Otherwise, the head of \( \gamma \) is not governed by \( \alpha \) externally. Let us define

---

2 The definition 3a is essentially the one proposed by Dominique Sportiche and Youssef Aoun. Belletti and Rizzi (1980) proposes the formulation:
   (i) The head of a maximal projection is accessible to an external governor but peripheral positions are not.
3 Chomsky (1986a: 162) defines the notion of govern as follows:
   (i) a category \( \alpha \) governs a maximal projection \( X' \) if \( \alpha \) and \( X' \) c-command each other; and if \( \alpha \) governs \( X' \) in this sense, then \( \alpha \) governs the Specifier and the head X of X'.
the condition on external government as in 4:

(4) Extraterritoriality (E.T.) Condition

External government is prevented in the structure:

\[ ... \alpha ... [\gamma ... \epsilon ... \delta ... \epsilon ...], \]

where \( \delta = X^0 \) and \( \gamma \) is a maximal projection of \( \delta \).

What is implied in 4 is that if there are not both an appropriate \( X^0 \) and \( \epsilon, \alpha \) externally governs \( \delta \). The existence of only one of them is not enough for \( \delta \) not to be governed by \( \alpha \) externally.

To illustrate the meaning of 4, consider the examples in 5, where arrows represent government relations:

(5) a. 

```
  VP  
 / \  
V   NP
  
destroy
```

b. 

```
  VP  
 / \  
  V CP(=COMP'=S')
 / \   
  C IP(=INFL''=S)
 / \  
    e NP  
    I'  
  I   VP
     [−tense]
   to
```

c. 

```
  CP  
 / \  
  C IP
  
for  
NP
  I'  
  VP
```

d. 

```
  VP  
 / \  
V PP
  
enter
  P
  into
  NP
```

e. 

```
  CP  
 / \  
  C IP
  
that
  NP  
  I'  
  VP
    [+tense]
    AGR
```

Following Chomsky 1986a and many other linguists, I will assume CP (=S') to be the maximal projection of COMP, IP the maximal projection of INFL and assume further that V, N, A, P, and Complementizer (e.g., for, that) and the agreement element AGR of INFL are the candidates for the appropriate \( X^0 \) in 4. In (a), V can govern NP and its head N, since in the NP there is no element which serves as a choice for \( \epsilon \) in 4. In (b), V can govern CP and further C, since \( e \) is not an appropriate \( X^0 \) for the choice of \( \delta \) in 4, if we assume that featureless \( e \) is “hopeless.”

in (c), C can govern INFL as well as IP, since to is not an appropriate choice for δ in 4.\(^5\) Note that the head N in 5a is externally governed because there is not any category which serves a choice for ε in 4, while the heads COMP and INFL in 5b–c are externally governed because those heads are inappropriate X\(^0\) for the choice of δ in 4. In case 5d, V governs PP but not P, since in the PP there are an appropriate X\(^0\) and NP which corresponds to ε in 4. In 5e, C (=that) can govern IP but not its head INFL with AGR, since IP contains both the appropriate candidates for δ and ε in 4, that is, AGR and NP.

Note that in the notion of government as in 3 and 4 we do not have to interpret the category of PRO in 2 as NP any more, but it is enough to regard the category of PRO as N when we say, “PRO is ungoverned.” If α governs \([_{\text{NP}} \text{PRO}]\), α also governs \([_{\text{N}} \text{PRO}]\) of the NP and if α governs \([_{\text{NP}} \text{PRO}]\) of \([_{\text{NP}} \text{PRO}]\), α also governs \([_{\text{NP}} \text{PRO}]\) as is clear from the consideration of the examples 5a and 5b.

Assume that PRO is lexically N (cf. Chomsky (1981: 162, 164)). Then PRO is a candidate for α in 3 as well as a candidate for δ in 4. Therefore PRO is a possible governor. But PRO is ungoverned. PRO is “selfish” in that it can govern governees (if there are governees) and may itself appear only in ungoverned positions in English.

3. Headless NPs. This notion of government has a rather interesting consequence in connection with the distribution of PRO (of N category). Consider examples 6 and 7:

(6) a. Distinguishing \([α \text{ meaningful}]\) from meaningless sentences is not always easy.
   b. Your preference of \([α \text{ tall}]\) to short girls is clear.

(7) a. Distinguishing \([_{\text{NP}} \text{meaningful sentences}]\) from meaningless sentences is not always easy.
   b. What dogs can you distinguish \(t\) from friendly dogs?
   c. \([_{\text{NP}} \text{Vicious dogs}]\) can be distinguished \(t\) from friendly dogs.
   d. Your preference of \([_{\text{NP}} \text{tall girls}]\) to short girls is clear.

Examples in 6 illustrate the phenomenon of what I call headless NPs. The objects of distinguish and of apparently lack their heads.

\(^5\) We assume to will get a governing capacity if it is externally governed, since it has some features. In 5c, NP is governed by to and for. But e in 5b will not get the capacity because it is featureless.
Let us examine the category labeled $\alpha$ in 6. That the lexical item *distinguish* in 6a takes NP as its object is clear from the examples 7a–c, where *distinguish* clearly takes NP as its object at some level of derivation. In 6b, the presence of *of* between the word *preference* and $\alpha$ implies that $\alpha$ is NP, on the assumption that the *of*-insertion rule takes place under the context [... N __ NP ...] or on the assumption that *of* generally takes NP as its object. The above support that $\alpha$ in 6 is NP.

If $\alpha$ in 6 is NP, then the head of $\alpha$ must be N from X-bar theory. Suppose that empty Ns in English are PRO and trace. Then the head of $\alpha$ in 6 is PRO, since there is no movement there and therefore no trace. Consider the partial structure of 6a and 6b as 8 and 9, respectively:

$$
\begin{align*}
(8) & \quad \text{VP} \\
& \quad \text{NP} \\
& \quad \text{N}' \\
& \quad \text{AP} \\
& \quad \text{N} \\
& \quad \text{distinguish meaningful PRO}
\end{align*}
$$

$$(9) \quad \text{PP} \\
\quad \text{NP} \\
\quad \text{N}' \\
\quad \text{AP} \\
\quad \text{N} \quad \text{I} \\
\quad \text{of} \quad \text{old} \quad \text{PRO}
$$

Note that these structures are exactly the ones which permit PRO to appear, that is, the positions in which PRO appears are ungoverned. In 8, *distinguish* governs NP but not PRO because of the E.T. condition (4): *distinguish* does not externally govern PRO, since AP and PRO are appropriate choices for $\varepsilon$ and $\delta$ in 4, respectively. Similarly in 9, *of* governs NP but not PRO. The same thing can be said to the following cases 10a–d:

---

6 One of the anonymous *English Linguistics* reviewers suggests that there is a possibility that the phenomenon of headless NPs involves similar mechanisms found in what is called Right Node Raising. But I will not pursue this possibility here.

7 In connection with the definitions 3 and 4, I am assuming here, following Jean-Roger Vernaud (1980) and Chomsky (1980, 1981), the Case filter of the sort shown in (i):

$$(i) \quad *\text{NP}, \text{where NP has a phonetic matrix but no Case (Chomsky 1981: 175)}$$

What (i) says is that “Every phonetically realized NP must be assigned Case” (Chomsky 1986a: 74).

Even if Case is assigned under government, the problem of whether the head N is governed or not is irrelevant to the Case filter, since the filter mentions the category NP, not the category N.

Consider the next examples, where all the head Ns in square brackets are ungoverned:
(10) a. Independence of local PRO from central government in this country is imminent.
b. Reliance of old PRO on young people is clear from the government report.
c. The similarity of the younger PRO to the older generation is getting less and less.
d. His theory gives an account of the annihilation of weak PRO by strong individuals.

The phenomenon of headless NPs is very interesting because the possibility of this phenomenon is what we would expect from the general principle: “PRO is ungoverned.”

It seems appropriate to ask the following questions as 11 when we find a marginal but quite interesting phenomenon, just as Chomsky (1982: 39) did when he treated parasitic gaps.

(11) a. Why does the phenomenon of headless NPs exist?
b. What are the basic properties of headless NPs?

Let us start with 11a. I have been assuming that PRO is N with some features. I will assume, following Chomsky 1982, that PRO is a pronoun without phonological content. Suppose that Universal Grammar permits choice of phonological content to be optional for choosing pronouns. Then an answer can be provided for the question 11a. Consider the list of English pronouns (12) and the list of subcategorization frames for N (13):

(12) a. I, you, she, he, it, we, they, one
b. PRO

(13) a. +[NP ————]
b. +[NP Specifier ————]  
c. +[NP Specifier ———— Complement]

(ii) a. ... destroy [the city]
b. ... destroy [a picture of Mary]
c. ... prefer [pretty PRO] to beautiful girls

Phonetically realized Ns like city and picture are not governed according to the definitions of 3 and 4: the and city in (iia) are appropriate candidates for $e$ and $\delta$ in 4, respectively, and a/of Mary and picture of (iib) correspond to $e$ and $\delta$ in 4 as well. But the whole NPs the city and a picture of Mary are governed by definition and are both assigned objective Case (cf. Chomsky 1981: 170). Therefore they are not ruled out by the Case filter (i).

In the same way the NP [pretty PRO] is governed, and it receives Case, satisfying the Case filter (i).
The members of 12a are pronouns with phonological content. All of them have the subcategorization frame 13a. The important thing here is that only the pronoun *one* has the frames 13b and 13c in addition to 13a:8

(14)  
\[
\begin{align*}
\text{a. } [\text{NP } [\text{One}]] & \text{ should feed oneself.} \\
\text{b. } & \text{This is my hat and that is } [\text{NP } \text{my brother's } [\text{one}]]. \\
\text{c. } & \text{Look at the pictures of Fred and } [\text{NP } \text{the } [\text{ones} ] \text{ of John}].
\end{align*}
\]

What is more important is that PRO can be a pronoun lacking phonological features of *one* on the assumption that PRO is a pronoun without phonological content. Therefore we have the examples as 15:

(15)  
\[
\begin{align*}
\text{a. } & \text{They told me how } [\text{NP } [\text{PRO}]] \text{ to feed oneself.} \\
\text{b. } & \text{This is my hat and that is } [\text{NP } \text{my brother's } [\text{PRO}]]. \\
\text{c. } & \text{Bill's funny story about Sue and } [\text{NP } \text{Max's } [\text{PRO}]] \text{ about Kathy] both amazed me.}
\end{align*}
\]

The answer for 11a is that a headless NP appears because of the optionality of phonological content for the pronoun *one*.

Turning to the question 11b, the basic property of headless NPs is more complicated than we would expect. Consider the next examples:

(16) I like  
\[
\begin{align*}
\text{a. } & \text{[Tom's } \text{PRO}] \\
\text{b. } & \text{[the rich } \text{PRO}] \\
\text{c. } & \text{*[friendly } \text{PRO}] \\
\text{d. } & \text{*[beautiful } \text{PRO}] \\
\end{align*}
\]

Although PRO in each example is not governed, (c) and (d) are not acceptable. Then why is there the grammaticality difference between 16c–d and 6a–b, 10a–d? (We can exclude examples like 16a–b from consideration of headless NPs, since NPs like *Tom’s* and *the poor* appear in typical NP positions without many restrictions.) We want to know

---

8 I am assuming that (1) the pronoun *one* falls in the subclass of the category N (i.e., pronoun) in the lexicon of English and (2) the pronoun *one* in each example of 14 differs in subcategorization but is the same lexical item. A question about the naturalness of the second assumption might arise: why does the pronoun *one* in 14a which is NP mean 'people' while *one* in the phrase *three black boxes and two white ones* for example, does not represent NP but N 'meaning 'box'?' This question might be answered in the similar way as might be done in the case of the transitive and intransitive verb pairs like *break the vase/*break. It might be possible to say that both of the transitive and intransitive verbs of *break* belong to the same lexical item *break* even if the subcategorization features of them are different.

The problem of whether *one* can have arbitrary reference or not is related with the internal structure of the NP in which the N *one* occupies the head position. As for the mechanisms of interpreting NPs with a pronominal head, see the explanation of 27 of this paper.
where headless NPs may appear. Look at further examples:\(^9\)

(17) a. To divorce logical from pragmatic aspects of meaning in an area of grammar ... (Hawkins 1978: 11)

b. derive passive from active sentences/convert active into passive sentences/turn active into passive sentences

(18) a. A generalization is proposed which unites semantic with syntactic causes of ungrammaticality.

(Hawkins 1978: 11)

b. comparison of legal to illegal acts/relation of legal to illegal acts

On the basis of the observation of the examples in 6, 10, 17 and 18, the basic properties of headless NPs can be stated as in 19 (and this is a first approximation):

(19) (i) a headless NP must be one of the logically necessary terms of the verb (or the noun),

(ii) those terms must be "SIMILAR" in a certain sense,

(iii) those terms must be in the complement positions of the verb/noun,

(iv) a headless NP must precede the other NP.

It can be said that if the noun phrases are indispensable in a certain sense to the action or the state denoted by the verb/noun, those NPs are called logically necessary terms. What 19ii implies is that if two terms are similar, they refer distinct entities and these entities share certain properties.\(^{10}\) Therefore the statement 19ii accounts for the unacceptability of the examples in 20:

(20) a. *I took small from many pebbles.

b. *a choice of far-advanced among senior students

In 20a, "small pebbles" and "many pebbles" cannot be interpreted as being distinct. Therefore we have the impossible 20a. The same is true of 20b.

The statement 19iii is about structural condition. Consider the examples in 21:

(21) a. We can distinguish friendly PRO from vicious dogs.

b. Friendly dogs can be distinguished t from vicious dogs.

---

\(^9\) The examples 17b, 18b and 20b are taken from Fukuchi (1985).

\(^{10}\) Postal (1971: 198) says, "To say that \(x\) is similar to \(y\) is to presuppose that \(x\) and \(y\) designate distinct entities and to assert that these entities share certain properties."
c. *Friendly PRO can be distinguished from vicious dogs.

d. *Vicious dogs can be distinguished friendly PRO from.

Compare (a) with (c-d). In (a), both terms of distinguish locate in the complement positions of the verb, while the headless NP in (c) and the NP vicious dogs in (d) are not in the complement positions after the movements of the NPs from the positions indicated as ti and tj, respectively. The structural condition 19iii accounts for the grammaticality difference between 21a and 21c-d.

Consider such examples as 22 and 23:

(22) a. derivation of passive sentences from active sentences
    b. derivation of [passive PRO] from active sentences
    c. passive sentences' derivation from active sentences
    d. *[passive PRO] derivation from active sentences

(23) a. dependence of old people on young people
    b. dependence of [old PRO] on young people
    c. old people's dependence on young people
    d. *[old PRO] dependence on young people

Compare 22b with 22d, and 23b with 23d. In 22b, the headless NP is on the right side of the derived nominal derivation, whereas in 22d it is not. The structural condition predicts the impossibility of 22d. The same holds true of the examples in 23.

The statement 19iv is also about structural condition. This condition accounts for the ungrammaticality of the next examples, in which the full NP precedes the headless NP:

11 The condition 19iv may be too strong as it is. An anonymous reader of English Linguistics points out that there are sentences like (i):

(i ) a. I prefer red wine to white PRO.
b. *I prefer red wine to white one.

But there seem to be more factors involved in the type of examples in which a full NP precedes a headless NP. It seems that one such factor is the countability of the head noun. One of the informants judges (iiia) as ungrammatical and (iiib) as grammatical:

(ii ) a. *I prefer tall girls to short PRO.
b. I prefer tall girls to short ones.

It seems that the head of a full NP must have the feature [+uncountable] when a headless NP may follow the full NP, whereas there is no such restriction when a headless NP precedes the full NP.

We exclude the examples of coordination like 15c from present consideration of headless-NP constructions.

Factors such as the kinds of specifiers, the types of adjectives, and the presence of stress may affect the acceptability of headless NPs, but I will not discuss them here, leaving them for future study.
(24) a. *Distinguishing meaningful sentences from [meaningless PRO] is not always easy.
b. *Independence of local government from [central PRO] in this country is imminent.

These arguments show that we need structural conditions of the sort shown in 19iii, iv in addition to “semantic” conditions as in 19i, ii in order to predict the distribution of the headless NPs correctly.

Now let us consider the implications of the restrictions 19. One might pose a question like 25:

(25) Why are there such restrictions as 19?

Before tackling this question, I should point out certain characteristics of “head PRO” (henceforth, H-PRO), that is, [N PRO] or [N PRO]. I will assume, following Yasui and Nakamura (1985: 5, 59), that only maximal projections have the “referential functions.” Therefore I will make a natural assumption like 26:

(26) H-PRO cannot have such a referential function as the usual NPs have.

Instead, it seems that H-PRO may take or borrow a “concept” from another word(s). Consider the typical example:

(27) Reliance of [NP old [N PRO]] on [NP young [N people]] is clear from the government report.

PRO in 27 must be interpreted as ‘people’ in order for the sentence to be understood correctly. Let us call this type of concept-borrowing as concept-control. Then we can say that people in 27 concept-controls PRO. In order for the NP [old PRO] to have a referent, the concept (or meaning) of PRO must be fixed first. The characteristic of H-PRO may be stated as follows:

(28) H-PRO must be in the position which guarantees concept-control in a certain way.

Compare the examples in 29:

(29) a. As observed first by Paul Postal, Condition (B) extends, though somewhat weakly, from distinct to disjoint reference, ... (Chomsky 1986a: 215)
b. Another familiar idealization is the assumption that the transition from initial to steady state can be regarded as instantaneous, ...
c. Choice of himself for $\beta$ ranges from dubious to impossible. (Chomsky 1986a: 129)
The partial structures of 29 in our analysis can be shown as in 30:

(30) a. Condition (B) extends, ..., from [distinct PRO] to [disjoint reference].
    b. the transition from [initial PRO] to [steady state]
    c. Choice of himself for β ranges from [dubious PRO] to [impossible PRO].

The examples are very interesting when we try to understand the meaning of H-PRO. PRO in (a) is interpreted as ‘reference’, PRO in (b) as ‘state’. PRO in (c) seems to be interpreted as something that is outside of the sentence. Note that even in the latter case H-PROs are exactly in the positions which are described in 19iii, iv.

A semantic characteristic of H-PRO is that H-PRO has a function of copying or “borrowing” the semantic features which constitute a concept like girl, for example. We can describe this property as follows:

(31) H-PRO must copy a concept expressed by other word(s).

First, H-PRO copies a concept of other word. Then the headless NP (which consists of the specifier and H-PRO) comes to have a referential function.

Let us return to the question 25. It is not unbelievable that since conditions of H-PRO like 28 and 31 must be satisfied, the distribution of headless NPs is severely constrained as described in 19. Suppose that H-PRO of a headless NP is concept-controlled by the head of the other NP. Then the condition 19i is interpreted as a “helpful” device for H-PRO in that it guarantees the existence of the concept-controller of H-PRO and the condition 28 is satisfied.

Suppose that H-PRO of one of the terms of a predicate must copy a concept of the head of the other NP. Then we can regard the condition 19ii again as a helpful device for H-PRO, in that if two terms are similar at a certain level, H-PRO can easily copy a concept expressed by the head of the other term. Then it follows that the condition of H-PRO (31) is satisfied.

Consider the condition 19iii again. The implication of 19iii seems to be that H-PRO needs the help of the predicate (or its nominal form) because it lacks phonological features. Since a verb, for example, carries the information about the subcategorization or about the number of terms, it is easier for speakers or hearers to expect the NPs to appear in the appropriate positions if the verb appears initially. It might be related to the manner of perception that the headless NP which lacks phonological features of the head is licensed as NP if it appears after the verb/
noun.

The condition 19iv also seems to be related to the PRO’s inherent property that it lacks phonological features. It might be reasonable that since headless NPs are “phonologically imperfect” NPs, they must occur in the positions which can be easily associated with some grammatical functions. NP-traces, which also lack phonological content, obey the similar restriction. Consider the fact exemplified by the next examples:

(32) a. Meaningful sentences can be distinguished _ from meaningless sentences.
   b. *Meaningless sentences can be distinguished meaningful sentences from _.
(33) a. old people’s reliance _ on young people
   b. *young people’s reliance of old people on _.

The positions from which NPs can be moved are exactly the positions where the headless NPs may occur. The headless NPs in these positions seem to function easily as NP, just as the NP-traces in the positions are easily associated with the full NPs.

The contrast between the examples in 34 can be evidence which supports the assumption that H-PRO is a pronoun without phonological features which corresponds to the pronoun one:

(34) a. Distinguishing [NP meaningful PRO] from meaningless sentences is not always easy.
   b. *Distinguishing [NP meaningful ones] from meaningless sentences is not always easy.

If we extend the principle 35 developed in Chomsky (1981: 65) to non-maximal categories, the grammaticality difference might be explained:

(35) Avoid Pronoun
The principle 35 imposes “a choice of PRO over an overt pronoun where possible.” Since H-PRO may occur in the position that the pronoun one occupies, H-PRO is chosen in that position and the example 34b which chooses the pronoun ones is ungrammatical. If the principle 35 is really at work in the examples in 34, the fact exemplified by 34 is a strong indication that H-PRO is a pronoun without overt phonological features.

4. HEADLESS NPs IN JAPANESE. It would be reasonable to investigate the phenomenon of headless NPs in other languages. Let me point out that the phenomenon also exists in Japanese, though it seems very marginal. Consider the examples in 36:
There are two interesting points to be made explicit. First, the headless NP in 36 cannot occur as the object of *kekkonsuru*, but it can occur as one of the objects of *konomu*. A similar observation is true of the examples in 37. The object NPs appear before the verb in each example. In English the NPs must appear after the verb (or the noun), as we have seen. Although we cannot have the exact formulation until we investigate many other languages, the possibility is that in head-first languages like English a headless NP appears after the head (V or N), whereas in head-final languages like Japanese it appears before the head.

Compare 36b and 37b: a second interesting point is that the order between the headless NP and the other NP seems to be rather free in Japanese. In 36b the headless NP comes first and then the full NP follows. In 37b the headless NP follows the full NP. But there seems to be an additional restriction on the choice of headless NPs with respect to the kinds of postpositions. Consider the next examples:


sukida.
The postpositions *ga and *o seem to hate headless NPs (cf. 36 and 37) while *yori seems to love them.
The factors which cause the differences between English and Japanese might be very interesting, but must be put aside for future research.
The phenomenon of headless NPs is difficult to find, but is worth studying, since it seems to be deeply related to some universal mechanisms, as we have seen.

5. Conclusion. Recently Chomsky 1986b tries to explain the distribution of traces uniformly, whether they are traces of maximal projections or traces of zero-level categories. Although he does not especially argue the distribution of PRO, it is reasonable to see the implications of his formulation of government which includes “minimality condition.” He formulates the condition as follows:

(39) In the configuration,

$\alpha \ldots [\gamma \ldots \delta \ldots \beta \ldots]$, $\alpha$ does not govern $\beta$ if $\gamma$ is a projection of $\delta$ excluding $\alpha$.
What 39 says is that $\delta$ prevents $\alpha$ from governing $\beta$. His formulation of “barrier” for the theory of government is 40:

(40) $\gamma$ is a barrier for $\beta$ if $\gamma$ is (a projection, the immediate projection) of $\delta$, a zero-level category distinct from $\beta$.
What 40 means is that in the configuration of 39 $\alpha$ cannot govern $\beta$ because $\gamma$ is a barrier for government. Notice, however, that $\gamma$ is a barrier for $\beta$, but 40 does not say $\gamma$ is a barrier for $\delta$, a zero-level category. This is a crucial difference between the E.T. Condition (4) and the minimality condition (39).
If it is the optimal assumption that the principles which apply to movement of maximal projections also apply to movement of zero-level categories, then it is reasonable to assume that the principle that applies to [NP PRO] applies to H-PRO, that is, zero-level PRO, or nonmaximal PRO.
In summary, the phenomenon of headless NPs can be seen as a consequence of this assumption; it can be argued that the headless NP phenomenon exists because phonological features are optional for the pronominal *one; moreover, the basic properties of headless NP, as in 19, have been clarified. I have speculated that these properties are highly related
to the H-PRO’s inherent characteristics: H-PRO needs a concept, it must be concept-controlled, and it lacks phonological features. I also pointed out that the phenomenon seems to be observed in Japanese as well.

Government for PRO is as stated in 3 and 4, while government by PRO is possible, and government of PRO is banned.

REFERENCES


