This paper offers an analysis of English infinitival relatives based on the framework of Chomsky's (1995) minimalist program. Infinitival relative clauses are analyzed as PPs headed by a null P, which take a CP as complement. We argue that this null head P has the property of attracting (an element with) the feature [+P(reposition)]. We show that various peculiar properties of infinitival relatives can be explained elegantly under our analysis from the interaction of general constraints such as the Head Movement Constraint, and a condition on the specifier-head agreement of the relative CP.

1. Introduction

English infinitival relative clauses (henceforth IRs) such as those in (1) are well-known for their peculiar properties.

(1) a. John bought a book [to read].
   b. John bought a book [for Bill to read].

Generally, overt wh-relative pronouns cannot appear in IRs.

(2) a. *John bought a book [which to read].
   b. *John bought a book [which for Bill to read].

However, an overt wh-relative can appear when a preposition is pied-

* I am deeply indebted to Jun Abe, Kinsuke Hasegawa, Yasuo Ishii, Howard Lasnik, Heizo Nakajima, Ken-ichi Takami, Kensuke Takonai, and two anonymous reviewers for insightful comments and suggestions on earlier versions of this paper. I am extremely grateful to Chris Tancredi for acting as an informant, giving me comments, and suggesting stylistic improvements. An earlier version of this paper was read at the TACL meeting held at Meiji Gakuin University on October 11, 1997. I have benefited greatly from discussion with the participants, among whom were Masaki Sano, and Shigeo Tonoike. I could not, however, incorporate all the comments and suggestions, due to limitations of time, space, and ability. All remaining errors and inadequacies, naturally, are my own.

English Linguistics 15 (1998) 1-27
© 1998 by the English Linguistic Society of Japan
piped in front of it, and for and an overt subject are absent. It cannot appear when the preposition is stranded, or when for and an overt subject are present.

(3) a. John bought a pen [with which to write].
   b. *John bought a pen [with which for Bill to write].

(4) a. John bought a pen [(for Bill) to write with].
   b. *John bought a pen [which to write with].
   c. *John bought a pen [which for Bill to write with].

These peculiar properties of IRs have so far defied a satisfactory analysis. Van Riemsdijk and Williams (1986: 168-170) make the following comment on data such as (2a):

(5) Several possibilities (none very attractive) exist to account for this fact (\(\cong(2a)\)) ....

After examining the problems of Chomsky and Lasnik’s (1977) analysis of IRs, they finish up the discussion by stating that a sentence like (2a) “remains recalcitrant.” McCawley (1988a: 453), after pointing out the problems of Emonds’ (1976) analysis of IRs, makes the following statement:

(6) I know of no alternative way to exclude such expressions (\(\cong(2a), (3b)\)) ... other than by brute force, e.g., a surface constraint ....

2. Previous Analyses

Before going into our analyses of IRs, we will examine previous analyses and point out their problems.

2.1. Emonds’ (1985) Analysis

Emonds makes somewhat different analyses of IRs in Emonds (1976) and in Emonds (1985), and we will briefly review the latter. According to Emonds (1985), the IRs in (1b) and in (3a) are derived as shown in (7) and (8), respectively:¹

¹ Emonds’ analysis is similar to ours in that IRs are PPs. In his analysis, however, every instance of the so-called S’ (including the IR) belongs to the category PP, and every C(OMP) is an instance of the category P. In our analysis, an IR is a PP which contains CP (=S’) as its complement.
We will not go into the details of his analysis, but it obviously involves a problematic operation. Both in (7) and in (8), a phrasal category (NP or P') is moved into a head (P) position. Such an operation is not permitted under the minimalist framework, and it is dubious whether it strictly conforms to the spirit of Emonds' Structure Preserving Constraint. Apart from such theoretical problems, his analysis also has an empirical problem, as pointed out by McCawley (1988a: 453).

(9) John bought a book for [[Bill to read] and [Mary to recite]].

As the grammaticality of (9) shows, the string "Bill to read" in (1b) forms a constituent. Under Emonds' derived structure in (7b) it does not form a constituent, and his analysis makes the wrong prediction that (9) should be ungrammatical.
2.2. Chomsky and Lasnik’s (1977) Analysis

Chomsky and Lasnik (1977) propose the following surface filter to rule out certain structures including certain IRs:

(10) Doubly Filled Comp Filter

\[
\ast \left[ \text{COMP } w_h\text{-phrase } \psi \right], \quad \psi \neq e
\]

The filter in (10) rules out the ungrammatical (11a), while allowing (11b):

(11) a. \( \ast \text{John bought a book } [\text{CP/S'} \text{ which_1 that } [\text{IP/S} \text{ he should read t}_i]] \).

b. John bought a book \([\text{CP/S'} \text{ which_1 } \psi [\text{IP/S} \text{ he should read t}_i]] \).

Even though we cannot resort to a ‘surface filter’ as a linguistic device in the minimalist framework, which has no S-structure as a linguistically relevant level, we will argue later that what amounts to this filter (see (28) below) is at work in ruling out certain ungrammatical structures, namely, IRs such as those in (2), (3b), and (4b, c) above.

Chomsky and Lasnik (1977) propose a more problematic filter:

(12) NP-to-VP Filter

\[
\ast [\alpha \text{ NP to VP}], \text{ unless } \alpha \text{ is adjacent to and in the domain of } [-N] (=a \text{ verb or } for).
\]

The \textit{unless} condition on the filter in (12) allows structures in the following configuration:

(13) \{ V \} \[\alpha \text{ NP to VP} \]

This \textit{unless} condition thus allows (1a) and (1b), due to the presence of the verb (bought) and for, respectively. On the other hand, (2a) does not have the configuration in (13), because a noun (book) is in front of \( \alpha \), and the sentence is ruled out by the filter in (12).

(2a') \( \ast \text{John bought a book } [\text{CP/S'}(=\alpha) \text{ which_1 (PRO) to read (t}_i)] \).

However, the filter in (12) incorrectly rules out (14), which has no V or for in front of \( \alpha \).

(14) \[\text{NP}(=\alpha) \text{ A book to read (t)} \] is on the table.

Thus, Chomsky and Lasnik (1977: 464) are forced to make the \textit{unless} condition on (12) more complicated:

(12') NP-to-VP Filter (revised)

\[
\ast [\alpha \text{ NP to VP}], \text{ unless } \alpha \text{ is adjacent to and in the domain of } [-N], \text{ or } \alpha = NP.
\]

Nevertheless, as Van Riemsdijk and Williams (1986: 170) point out, even this revised filter in (12') incorrectly rules out the grammatical sentence in (15):

(15) It is unclear \([\text{CP/S'}(=\alpha) \text{ what_1 (PRO) to do (t}_i)] \)
Sentence (15) is not exempt from the filter by the *unless* condition, for the adjective (*unclear*) in front of $\alpha$ is $[+N]$, and $\alpha \neq \text{NP}$. The NP-to-VP filter, even with its ad hoc *unless* condition, has its empirical flaws.

### 2.3. Chomsky's (1980) Analysis

Chomsky (1980) assumes the Case Filter, which accounts for most of the cases covered by the NP-to-VP Filter in Chomsky and Lasnik (1977).

(16) Case Filter

*\(N\), where N has no Case

However, the above Case Filter is not enough to explain all the data previously covered by the NP-to-VP Filter: for instance, it cannot handle the ungrammaticality of (2a). In order to account for such data, Chomsky (1980: 22) claims that deletion of elements in COMP is obligatory before infinitives, under the assumption that deletion of *which* is operative in the derivation of the grammatical (1a):

(17)  

As shown in (18), *which* moved to COMP is obligatorily deleted; only the grammatical (1a) is derived, and the ungrammatical (2a) is ruled out.

(18) John bought a book \([\text{CP/S'} \text{ which}_{i} \text{ [IP/S (PRO) to read } t_i\text{]}\].

Why then, is (3a) grammatical, even though some elements (*with which*) remain in COMP?

(3a') John bought a pen \([\text{CP/S'} \text{ [PP with [NP which]}_{i} \text{ [IP/S (PRO) to write } t_i\text{]}\].

Chomsky's (1980) account goes as follows. Deletion of the entire PP (*with which*) is impossible, because the deletion of the preposition (*with*) would violate the recoverability condition. He goes on to argue that we cannot delete only *which*, leaving *with* intact: hence, the elements in COMP “survive deletion” and (3a) is grammatical.²

² Apart from the dubious character of such “obligatory deletion” in grammar (which is probably one of the reasons why empty operators were introduced in
But why can't we delete only the *wh*-relative (*which*), leaving the pied-piped preposition (*with*) intact? Chomsky (1980: 21) claims that such deletion is impossible "for principled reasons," and refers the reader to Chomsky and Lasnik (1977: 446, fn. 43), where it is suggested that some form of the A-over-A Principle is operative. In (3a') above, however, *with which* is a PP while *which* is an NP, and this is not a typical A-over-A configuration. Moreover, there are empirical data which seem to suggest that the reason for the impossibility of such ‘partial deletion’ is not so "principled" as Chomsky (1980) seems to believe. According to Visser (1963–73: §930), IRs such as those in (19) were possible in Old and Middle English.

(19) a. a foot on to goo (a foot on which to go)
   b. a hous in to drink and ete (a house in which to drink and eat)

Under the assumption that IRs without overt *wh*-relatives are derived by deletion of *wh*-relatives, the IRs in (19) are exactly the cases where the *wh*-relatives (*which*) are deleted while the pied-piped prepositions are left intact. It seems difficult to come up with an exact formulation of the universal A-over-A Principle which accounts for both the impossibility of partial deletion in (3a’) and its possibility in (19).

Chomsky (1981) and in subsequent works), there seems to be a problem of how elements can “survive” deletion when the application of deletion is blocked, if such deletion is in fact “obligatory.” Chomsky (1980: 21) claims that “obligatory deletion” means “delete wherever possible,” and that its inapplicability enables the elements in COMP in (3a) to “survive obligatory deletion.” It is unclear why the inapplicability of an “obligatory” operation results in “survival” from the operation, rather than in ungrammaticality.

3 Chomsky and Lasnik (1977: 446) claim that the deleted *α* in COMP can be of “an arbitrary category,” and thus the deletion of the relative pronoun contained within a PP is excluded by A-over-A, even though the deleted element (NP) and the larger element including the deleted element (PP) are not of the same category.

As has been suggested by Masaki Sano (personal communication), it might be possible to rule out such ‘partial deletion’ by formulating the A-over-A Principle in terms of features such as [+wh], on the assumption that the [+wh] feature percolates up to the PP node in the case of pied-piping. See Chomsky (1977a: fn. 55) for a related discussion. See also Fukui (1996) for an attempt to reformulate the A-over-A Principle in terms of features.

4 If we pursue the possibility of the ‘feature relativized’ A-over-A Principle suggested in footnote 3, we would have to claim that the A-over-A Principle is somehow parameterized with respect to the [+wh] feature.
3. Our Analysis of the Infinitival Relative

3.1. The Infinitival Relative as a PP Headed by a Null P

Assuming Chomsky's (1995) framework of the minimalist program, we will make the following set of proposals as an analysis of IRs:

(20) a. An IR is a PP headed by a null P which takes a (infinitival) CP as its complement.
   b. A (head containing a) \([+P]\) feature must move(/be attracted) to the null head P.\(^5\)
   c. The complementizer *for* is a ‘prepositional complementizer’ which has a \([+P]\) feature, and it must move to the null head P position.
   d. The infinitival *to* (optionally) contains a \([+P]\) feature, and it moves to the null head P position (when *for* is absent), after the movement to the C position.

Our claim in (20a) seems natural since prepositions like *before* take clausal complements as in (21b):

(21) a. the day *before her arrival*
   b. the day *before [she arrived]*

In fact, Larson (1987) (see also Johnson (1988)) argues that *before* takes an \(S' (=CP)\) in structures like (21b), and that movement of an empty operator is involved, which is quite similar to our analysis of IRs. Koster and May (1982: fn. 24) give the following instance of a P taking an infinitival \(S' (=CP)\), within which an operator moves to (the Spec of) Comp:

(22) He asked me \([PP about [S'/CP who to visit (t)]]\)

It is often claimed that an adjunct infinitival clause such as the one in (23) (which is a ‘rationale clause’ in Faraci’s (1974) terms) is a PP (see also Emonds (1985)).

(23) Bill bought the piano \([PP (in order) (for Mary) to practice on it]]\)

It is natural that IRs, which are adjunct infinitival clauses of another type, should also be PPs.\(^6\) Since PPs often appear as post-noun modifiers as in (24) below, it is again not unnatural that an IR, which is a

\(^5\) We will argue later that this movement to the null P position is not overt categorial movement but actually covert movement of a \([+P]\) feature.

\(^6\) It seems plausible to speculate that all adjunct infinitival clauses are PPs, but we will not pursue this speculation further here, due to limitation of space.
kind of post-noun modifier, be analyzed as a PP:  

(24) a. a book [PP for children]  
b. the key [PP to success]

Jackendoff (1977: 187–190) observes that a PP precedes a finite relative clause when the two co-occur as post-modifiers of an NP:

(25) a. a man [PP with a scar] [who came to dinner]  
b. *a man [who came to dinner] [PP with a scar]

(26) a. a pan [to fry omelets in] [that is stainless]  
b. *a pan [that is stainless] [to fry omelets in]

As pointed out by Jones (1985: 21), an IR must also precede a finite relative clause when they co-occur, which suggests that an IR is a kind of PP.

A claim similar to our claim in (20c) that the complementizer for has prepositional properties has been made by many other linguists (cf. Emonds (1976, 1985), Chomsky (1980: 30), etc.). See Abe (1986) for an argument that the infinitival to may have prepositional properties, as we have claimed in (20d).

We speculate here that the null head P of an IR might be a syntactically defective P which cannot be interpreted as a full-blown preposition at LF. It would then have to attract (some other element with)

---

7 We do not intend to claim that all post-noun modifiers are PPs. For instance, there seems to be no evidence that finite relative clauses are PPs. As pointed out by Yasuo Ishii (personal communication), an IR can be modified by just or not as in (i), but a finite relative clause in (ib) cannot be.

(i) a. Here is a book [just | [for children to read].
    not

b. *Here is a book [just | [that children should read].
    not

c. Here is a book [just | [PP for children].
    not

Since an ordinary PP can be modified by just or not as in (ic), the above data (indirectly) suggest that an IR is a PP while a finite relative clause is not.

8 According to Fukui (1995), complementizers, which have the features [+F(unctional), −N, −V], are close in nature to prepositions, which are [+L(exical), −N, −V]. We might claim that the English complementizer for is even closer to P in that it is both +F and +L: Fukui (1995: fn. 3) suggests the possibility of such a feature combination.

9 Tanaka (1997) argues that the infinitive marker to, which was a P(reposition) in Old English, went through a categorial change and became a T(ense) in Middle English. We claim that the infinitival to may retain some of its prepositional properties, at least when it appears in IRs.

10 We might also speculate that the null head P of an IR has a semantic content
a categorial feature [+P] for LF convergence: hence the requirement in (20b). Technically, this would mean that the null head P of an IR in English has a -interpretable feature which must be checked off by attracting (some other element with) a +interpretable feature [+P].

Having made our proposals, let us see how they work. First, consider the derivation of sentence (1b) shown in (1b'):

(1b') John bought a book \([_{PP} \phi \ [_{CP} \phi_{1} \ 1_{C'} \ 2_{IP} \ \text{Bill to read } t_{i}]]\)

The complementizer for has the [+P] feature, and it moves to the null head P position; the requirement in (20b) is satisfied, and the sentence is grammatical. Now let us look at the derivation of (1a) in (1a'):

(1a') John bought a book \([_{PP} \phi_{1} \ 1_{CP} \phi_{2} \ 2_{IP} \ (PRO) \ to \ read \ t_{i}]\)

First, the infinitival to with the [+P] feature moves to the null head C position (indicated by \(\phi_{2}\)). This is followed by movement to the null P position (indicated by \(\phi_{1}\)), satisfying (20b).

A question arises here of what the ‘driving force’ of the first movement (the movement to null head \(\phi_{2}\)) is. We adopt Pesetsky’s (1995: 8) idea that a null complementizer is a null affix. We claim that this null affix \(\phi_{2}\) induces the syntactic movement of to to the head C position in (1a'). Informally speaking, an affix is an element that cannot ‘stand on its own’ and must be ‘associated with’ some other element. Such a requirement is formulated as a filter in Lasnik (1981), and in Baker (1988):

roughly equivalent to that of the preposition for; (ia) is roughly equal in meaning to (ib):

(i) a. a paper \([_{PP} \phi \ [(PRO) \ to \ publish \ (t)]]\)
   b. a paper \([_{PP} for \ [\text{publication}]]\)

See Browning (1987: 315–317) for a similar observation.

11 We assume movement of an empty operator, even though our argument here does not hinge crucially on this assumption.

12 As we will argue later, this (second) movement from the head C \(\phi_{2}\) position up to the head P \(\phi_{1}\) position is actually covert movement of the feature [+P], while the (first) movement to the C position is overt (categorial) movement of to.

13 We do not claim that null affixes always induce syntactic movement: see (32) below. In the case of (1a'), the infinitival to must move syntactically to the null head C position. Otherwise, the null head P could not attract (the [+P] feature
(27) Stray Affix Filter (Baker (1988:140))
\[*X if X is a lexical item whose morphological subcategorization frame is not satisfied (at S-structure).\]

Baker (1988) regards (27) as an S-structure filter. Under the minimalist framework where the level of S-structure is gone, it seems plausible to assume that such a requirement applies at PF: one cannot pronounce an affix if it is not associated with some other element, and the derivation crashes at PF.14

Next let us consider how the ungrammatical IRs in (2)-(4) are ruled out. As we have already suggested above, we will assume what amounts to Chomsky and Lasnik's (1977) 'Doubly Filled Comp Filter' in (10). Instead of a 'surface (S-structure) filter', however, we will propose a condition on Specifier-Head agreement, along the lines of Rizzi (1990: 66-68), informally stated in (28):

(28) If the Spec of (a relative) CP contains a wh-relative, the head C must be null.

The condition in (28), for instance, rules out (11a) and allows (11b), reproduced below:15

of) to, since the null head C (\(\phi_2\)) would act as an intervening head for the HMC. For comparison, see (2a'') and (3a'') below.

As pointed out by Heizo Nakajima (personal communication), we must explain how the PRO subject can have its Case checked, under the assumption that a PRO subject gets its 'null Case' checked by the infinitival to (cf. Chomsky (1995: 119-120)). One possibility is that the Case feature of PRO moves covertly to the raised infinitival to and is checked in the C position. Another possibility is that the null Case of PRO is checked prior to overt I-to-C movement (of to). Case checking of infinitival subjects before overt I-to-C movement is a possibility suggested in Watanabe (1996:61) for for-to infinitivals.

14 As pointed out by Chris Tancredi (personal communication), we must assume that (27), which we take to apply at PF, applies also to null affixes, even though null affixes are not actually pronounced.

15 We assume here that (28) applies only to relative CPs, which are headed by Cs with the feature [+Pred(icative)], according to Rizzi (1990). The sentences in (i) below are not ruled out as 'doubly filled comp' violations, since the head C of the bracketed CP is \([-\text{Pred}]\):

\[
\begin{align*}
(\text{i})\quad & \text{a. } [\text{CP } \text{What}_2 \text{ did } [\text{John buy } t_1]]? \\
& \text{b. } \text{What}_2 \text{ did Bill say } [\text{CP } t_1 \text{ that } [\text{John bought } t_1]]?
\end{align*}
\]

The condition in (28) might apply at LF, if it is an interface condition at all. However, if we assume, along the Lasnik and Saito (1992), that a complementizer such as that has no semantic content and can be deleted at LF, (28) interpreted as an interface condition might not be able to rule out at LF the 'doubly filled comp' violation in (11a). See (28') below for an alternative formulation based on mis-
ENGLISH INFINITIVAL RELATIVES AS PREPOSITIONAL PHRASES

(11) a. *John bought a book \[\text{CP which} \text{that IP he should read}_t \].

b. John bought a book \[\text{CP which} \text{IP he should read}_t \].

This condition rules out the ungrammatical IRs with ‘doubly filled comp’ in (2b), (3b), and (4c) as desired:

(2b') *a book \[\text{which for Bill to read}_t \]

(3b') *a pen \[\text{with which for Bill to write}_t \]

(4c') *a pen \[\text{which for Bill to write with}_t \]

The derivation of the problematic (2a) is indicated in (2a''):

(2a'') *John bought a book \[\text{PP } \text{[CP which} \text{IP PRO} (t) \text{read}_t \] \].

In order to satisfy the requirement in (20b), an element with a \[+P\] feature must move to the null P position. The only candidate in (2a'') is the infinitival to. However, it cannot move directly to the null P position, because the movement will violate the Head Movement Constraint (HMC) by skipping the intervening head (C) position. Therefore, to must first move to the head C position, which creates a ‘doubly filled comp’ configuration together with which in the spec of CP, and the structure is ruled out by the condition in (28). (The ungrammatical (4b) above is ruled out in the same manner.)

One might wonder why the derivation in (2a'') cannot escape violation of (28) by the movement of to in the C position to the null P above. One possible reason is that even after the movement of to up to the null P position, a trace (or a copy) of the moved element remains in the head C position. This trace (copy) of to might act as a non-null head C which creates a ‘doubly filled comp’ violation together with the \text{wh-relative} in the Spec position.

However, it is unclear whether (a trace or copy of) to should really match of \text{wh}-features.

16 If the movement to the null head P is the covert movement of a feature \([+P]\) as we will argue later, we will have to reformulate the HMC in terms of features. As we have seen in footnote 8, complementizers and prepositions have common features, and it seems plausible that these common features of C act as an ‘intervening head’ for the HMC with respect to the movement of a \([+P]\) feature.

17 If the movement to the null P position is in fact covert movement of the feature \([+P]\) as we will argue later, everything except the \([+P]\) feature of to will remain in the C position, which can potentially cause a ‘doubly filled comp’ violation.
remain in the C position at LF; in light of the principle of Full Interpretation, it should not if it is not interpreted in this position.\(^{18}\) Without to in the C position, (2a'') cannot be ruled out by (28) at LF as a ‘doubly filled comp’ violation.

Chomsky’s (1995: 309) claim in (29) below suggests another (perhaps more plausible) reason:

\[(29)\) Mismatch of features cancels the derivation.

Once the infinitival to moves into the head C position in (2a''), it causes some mismatch of features with the wh-relative (which) in the Spec position: the derivation is cancelled, and no further derivation is possible.

According to Rizzi (1990), one of the features relevant to this ‘doubly filled comp’ effect is \([\pm \text{wh}]\). Rizzi (1990) claims that an overt wh-relative pronoun is specified as \([\pm \text{wh}]\) while an empty relative operator is not specified with respect to \([\pm \text{wh}]\). We might assume that the null complementizer in a relative clause is neutral with respect to the feature \([\pm \text{wh}]\) and is compatible with either an overt wh-relative or an empty operator, while the complementizers that and for are specified as \([-\text{wh}]\) and are incompatible with overt wh-relatives. The condition in (28) can then be restated as in (28') in terms of the feature \([\pm \text{wh}]\):

\[(28')\] If the Spec of a relative CP contains an element with the feature \([+\text{wh}]\), the head C of the CP must not have the feature \([-\text{wh}]\).\(^{19}\)

If the relevant feature is \([\pm \text{wh}]\), we have to assume that to in an IR has the same feature as for, namely, \([-\text{wh}]\), which causes the cancellation of the derivation in (2a'') above. This is not an unwanted consequence, for a \([+\text{wh}]\) relative pronoun can never be raised from the

---

\(^{18}\) This has been pointed out by Howard Lasnik (personal communication).

\(^{19}\) Lasnik and Saito (1984: 240) (see also Lasnik and Saito (1992: 3)) propose a similar condition:

\[(i)\] *Comp, if it contains a \([+\text{wh}]\) element.

\([-\text{wh}]\)

An anonymous reviewer points out that a condition like (28') cannot rule out the following ungrammatical sentence:

\[(ii)\] *John asked \([\text{CP which} [\text{IP if Mary ate t}]]\].

\([+\text{wh}]\) \([-\text{wh}]\)

Though interrogative clauses are beyond the scope of this paper, it seems plausible to speculate that the above sentence is ruled out by its uninterpretability at LF, rather than as a ‘doubly filled comp’ violation.
Spec of TP (=Subject) position in an IR:\(^{20}\)
(30) *John found a man [(for) who(m) to fix the sink].

If this line of reasoning is correct, (28’) might be a requirement on possible spec-head relations that must be satisfied in the course of derivation, rather than an LF interface condition.\(^{21}\)

3.2. Overt Wh-Relative with a Pied-Piped Preposition

Now why is (3a) grammatical even with an overt wh-relative? The presence of a pied-piped preposition (\textit{with}) within the Spec of CP makes the difference:

(3a’’) John bought a pen [PP \(\phi\) [CP [PP with which\(i\) [C’ \(\phi\) [IP (PRO) to write \(t_i]\)]\] [+P]]

The preposition with has a [+P] feature which is attracted by the null head P above, and the requirement in (20b) is satisfied. This structure, schematized in (3a’’’) below, is not in a typical configuration of head movement; the movement originates from within the specifier position of CP:

(3a’’’)

\[
\begin{array}{c}
\text{PP} \\
\downarrow \\
P \\
\phi \\
\downarrow \\
\text{PP} \\
\downarrow \\
\text{CP} \\
\downarrow \\
C' \\
\downarrow \\
P \\
\text{NP} \\
\text{C} \\
\text{IP} \\
\text{with which} \\
\downarrow \\
\text{[+P]}
\end{array}
\]

\[
\begin{array}{c}
\text{VP} \\
\downarrow \\
V \\
\downarrow \\
\text{CP} \\
\downarrow \\
C' \\
\downarrow \\
\text{VP} \\
\text{C'}
\end{array}
\]

\(^{20}\) We do not claim that every instance of the infinitival \textit{to} in English has the feature \([-\text{wh}]\) which rules out a \([+\text{wh}]\) element in the Spec (\(=\) Subject) position. The following structure is not ruled out in the same manner as (30):

(\text{i}) Who, do you believe [IP \(t_i\) to be competent]?\(^{21}\)

\(^{21}\) This ‘cancellation by feature mismatch’ approach might account for the ungrammaticality of the following example from Ishii (1985):

(\text{i}) *I found a cot [on which\(i\) to arrange [\(\_\_\_\_\_\_\_\_\_\text{ for [Mary to sleep } t_i]\)].

\[\_\_\_\_\_\_\_\_\_\text{ for [Mary to sleep } t_i]\]
Baker (1988: 170) argues for the possibility of a similar head movement in (31) above, with a head V moving out of the VP in the specifier position of CP to a higher head position. We will follow Baker (1988) and assume that such a head movement is possible.

In order to rule out the ungrammatical (2a), we argued above that the attraction of the infinitival to to the head C position in (2a") creates a ‘doubly filled comp’ configuration, which violates (28) (or (28')). We claimed that this movement was driven by the ‘affixhood’ of the null head C. Why, then, isn’t (3a) ruled out in the same manner?

Adapting Bobaljik’s (1994, 1995: Ch. 2) idea of affixation, we claim that there are (at least) two possible ways that an affix can be ‘associated with’ some other element:22

(32) a. adjunction via head-to-head movement in syntax (followed by ‘morphological merger’, in Bobaljik’s terms)
   b. morphological merger with (i.e. phonetical realization on) a phonetically adjacent element

In other words, the ‘association’ of an affix with another element may either involve or not involve syntactic head movement. In the case of the grammatical (3a), the null affix C may be associated with a phonetically adjacent element in the ‘phonological component’: that is, after Spell-Out, before PF. No syntactic head movement of the infinitival to to the head C position is necessary; hence no ‘doubly filled comp’ violation.

Note that the ungrammatical (2a) cannot escape a ‘doubly filled comp’ violation in the same way as (3a). As we have already argued, the infinitival to, which has the [+P] feature, cannot be moved/attracted directly to the null P position above; the intervening head C causes an HMC violation. Therefore, in order to satisfy the requirement in (20b), the infinitival to with the [+P] feature must first move

Once the wh-relative moves into the intermediate spec of CP position, a ‘doubly filled comp’ configuration (due to the presence of for) creates a feature mismatch: no further movement upward to save the structure is possible, even though the C upstairs is null. We need further examination of various cases to see if this approach really goes through.

22 Strictly speaking, Bobaljik’s notion of ‘adjacency’ is not purely phonetic; for instance, two elements are ‘adjacent’ when the intervening element is an adverb, according to Bobaljik (1995: 57). Note also that Bobaljik (1995: §3.1) argues against regarding a morpho-phonological feature [+affix] as a trigger for syntactic movement.
to the head C position (in syntax), which results in a 'doubly filled comp' violation. In the case of the grammatical (3a), on the other hand, the preposition *with* within the Spec of CP has the [+P] feature, and it moves to the null P position to satisfy (20b). There is no need for the infinitival *to* to move to the head C position in syntax. The null affix C can be associated with a phonetically adjacent element (in the 'phonological component'), and satisfy its affixal requirements without inducing any syntactic movement.

4. Further Data and Analyses

In this section, we will show that our analysis provides for a straightforward account of certain IRs, which, as far as we know, have not been analyzed thoroughly in the previous literature.

4.1. Empty *Where* in IRs

First, compare the sentences in (33):

(33) a. John found a place \[[[NP Op_i] (for Mary) to | live in | [NP t_i]].
    \[go to\]

b. John found a place \[[[PP Op_i] (for Mary) | live | [PP t_i]].
    \[go\]

Since verbs such as *live/go* are subcategorized for PP complements, the trace after the verb in (33b) must be a trace of a PP. On the assumption that movement of an empty operator is involved in these IRs, the moved empty operator in (33b) must also belong to the category PP. It is not the case that this empty operator PP can replace any kind of PP. The example below is ungrammatical without the preposition *with*; the empty PP cannot replace a PP headed by *with*:

(4a') John bought a pen \[Op_i (for Bill) to write *(with) t_i].

We claim that the empty operator PP involved in (33b) is an 'empty where.' In fact, there are instances of 'empty when' and 'empty how'

---

23 The possibility of an empty operator that belongs to the category PP is suggested in Johnson (1988: fn. 6). Larson (1985) claims that 'bare-NP adverbs,' including the relative *when* and *where*, are not PPs but NPs. We follow Emonds (1987) and McCawley (1988b) and assume that what Larson calls 'bare-NP adverbs' are PPs.
in IRs, illustrated in (34):^{24}

(34)  
\[ \text{a. John told Mary the time [Op; to leave t].} \]
\[ \text{b. John told Mary the way [Op; to do it t].} \]

The 'empty when' in (34a) amounts to "phonologically null when" argued for in Larson (1987) and Johnson (1988) in their analyses of before clauses. We will concentrate on 'empty where' here.

Even though either an overt where or an 'empty where' is possible as a relative pronoun in an ordinary relative clause, as in (35a), an overt where cannot appear in an IR like (35b); only an 'empty where' is allowed. On the other hand, an overt relative (which) with a pied-piped preposition can appear in an IR, as in (35c):

(35)  
\[ \text{a. John found a place [CP [PP [where i [C (\phi) [IP Mary could live/go [PP t]]]]].} \]
\[ \text{b. John found a place [PP \# [CP [PP [\*where i [C to [IP (PRO) (t) [+P]]]]]]].} \]
\[ \text{c. John found a place [PP \# [CP [PP [in/to which i [C (\phi) [+P]]]]].} \]

The contrast in (35) follows directly from our analysis. In (35a), neither the overt where nor the 'empty where' violates (28) (or (28')), with a null C as the head. In (35b), the infinitival to (which has the [+P] feature) must first move syntactically to the head C position before the movement up to the null P position, in order to satisfy the requirement in (20b).^{25} The overt where creates a 'doubly filled comp' configura-

^{24} McCawley (1991: 51) points out the following example from Shakespeare:

( i ) ... but teachest me the way how to lament the cause. (Richard II)

Since overt wh-relatives in IRs are ruled out as 'doubly filled comp' violations in our analysis, the possibility of an overt how above might be related to the possibility of the sequence how that in English of this period. We will not elaborate on this further here.

^{25} We are assuming here that the overt relative where cannot serve in any way to satisfy the requirement in (20b). It might be argued that since where can be analyzed as a PP, it might also be regarded as an 'intransitive' head P with a [+P] feature, under the minimalist framework where a category can be both minimal and maximal at the same time. Even so, where cannot serve to satisfy (20b); since the
tion, and is ruled out by (28'), as in the case of (2a''). As pointed out by Jun Abe and by an anonymous reviewer, there might be a possibility of analyzing IRs as CPs, the head C of which requires a [+P] element either in the Spec position or in the head position. Under this analysis, (1b) would be an instance of a [+P] element (for) in the head C position, while (3a) (or (35c)) would be an instance of a [+P] element (PP) in the Spec of CP position. This analysis, however, cannot account for the ungrammaticality of (35b) with an overt where in the Spec position, which seems to have the properties and distribution of a PP (cf. fn. 23). In our analysis, the null head P of an IR attracts the [+P] feature of a 'head', and only a 'headed' PP in the Spec position can save the structure from a 'doubly filled comp' violation (as in (3a) and (35c)). (See also footnote 25 for a comment related to the 'headedness' of where.) The acceptability of (36) and (42) below might also pose a problem for such an analysis. Since the element in the Spec of CP in these sentences is not a PP but an NP as a whole, these sentences are predicted to be ruled out on a par with an ungrammatical sentence like (2a) under this analysis.

27 'Doubly filled comp' configurations in sentences below are also ruled out by (28) (or (28')), in the same way as (2b) or (3b):

(1) a. *John found a place [where for [Mary to live/go t]].
   b. *John found a place [in/to which for [Mary to live/go t]].

28 Ishii (1985) gives the following example as an unacceptable one:

(1) "He bought a book [[the content of which] to discuss (t) with Mary]. According to my informants, however, IRs such as those in (36) are (marginally) acceptable, or at least better than an IR like (2a).
IRs like those in (36) seem to be clearly better than ill-formed IRs with overt *wh*-relatives like (2a). The acceptability of (3a) (in which the overt *wh*-relative co-occurs with a preposition) indicates that the italicized preposition (*of*) within the Spec of CP makes the difference. Since there is no separate overt movement of the preposition (*of*) here, the acceptability of (36) suggests that the movement to the null head P position is actually covert movement of the feature [+P]. In other words, the null head P of an IR has a 'weak' feature, which induces 'covert head movement' of the feature [+P], as shown in (36') below:

(36')

The infinitival *to* need not move up to the head C position in syntax, as was the case in (3a'') above; hence no ‘doubly filled comp’ violation.29

The following examples constitute a piece of evidence that this movement up to the null head P position is in fact covert feature movement, not overt (categorial) movement:

(37)  a. *a pen [PP with [CP (t) which to write] and [CP (t) which to poke her in the back]]

29 Sentence (i) below is again ruled out by (28'):

(i) *The carpenter is looking for a house [the roof of which for [his son to fix t]].
ENGLISH INFINITIVAL RELATIVES AS PREPOSITIONAL PHRASES

b. a pen \([PP \neq [CP \text{ with which to write}] \text{ and } [CP \text{ with which to poke her in the back}]\)

If the movement to the null head P position were overt categorial movement, the preposition \((with)\) should be able to undergo overt 'across-the-board' movement as in (37a). However, (37a) is ill-formed, which indicates that the movement to the null P is not overt head movement (of with). The acceptability of (37b) suggests that the movement to the null head P position is actually covert movement of the feature \([+P]\), which in (37b) is an 'across-the-board' movement.

We will now speculate why IRs like those in (36) are less acceptable than IRs like (3a), even though the two types of IRs both escape a 'doubly filled comp' violation due to the presence of a preposition within the Spec of CP. The marginality of IRs of the type in (36) might be related to the fact that the movement to the null head P is an extraction out of NP in these cases. Extraction out of an NP often results in lower acceptability, which Horn's NP constraint tried to capture (cf. Bach and Horn (1976:280)):

(38) No constituent that is dominated by NP can be moved (or deleted) from that NP by a transformational rule.

However, the difficulty of extraction from NP varies from case to case, as argued by many linguists (cf. Chomsky (1977a, 1977b), Bach and Horn (1976), etc.). Moreover, it is not clear whether such a constraint on movement applies uniformly to covert feature movement also.

Another possible way of explaining the degraded acceptability of (36) is that the head noun \((cover\) in (36a)) might act as an 'intervening head' for the HMC when the \([+P]\) feature (of of) moves up to the null P position. But if it is a true violation of the HMC, why shouldn't sentences like (36) be totally ungrammatical, rather than just marginal? Safir (1986) argues for the 'LF movement in comp' of relative pronouns in relative clauses schematized in (39) below:

(39) Ni \([S'/CP [wh-rel [NP ... e_i]] j [S/IP ... e_j ...]]\)

If such LF (covert) movement can in fact take place, it is possible that a preposition may be pied piped along with the moved relative pronoun, as shown in (40) below:\(^{30}\)

---

\(^{30}\) See Hirata (1997) for an analysis of relative clauses in a similar configuration in
(40) a book $[\text{PP } \phi \ [\text{CP } \{	ext{PP of which}, [\text{NP the cover } t_i], [(\text{PRO}) \text{ to } ] + \text{P} \}\]]$

The preposition (of) may thus skip the intervening head noun (cover), and escape an HMC violation. The marginality of (36), then, may be caused by one of the following two factors: (I) the LF moved wh-relative (which, which is an operator) does not strictly c-command the trace within (the Spec of) Comp ($t_i$, which should be a variable bound by an operator) because of the presence of the PP node, or, (II) the movement of the [+P] feature up to the null P position originates from this LF moved element (PP), which is perhaps in an adjoined position; movement out of an adjoined element generally results in lower acceptability, but not always in total ungrammaticality. Bošković (1997: 26) (see also Lasnik and Saito (1992: 101)) gives the following example, in which extraction out of an adjoined element results in marginality rather than in total ungrammaticality:

(41) ?Who do you think that [friends of $t_i$], [Mary kissed $t_i$]? We will leave further clarification of these problems for future research.

4.3. Possessive Relative Pronouns in IRs

According to McCawley (1988a: 429–454, 1988b: fn. 3), some speakers “accept and spontaneously produce” IRs of the following sort, in which possessive wh-relatives appear overtly:\footnote{Emonds (1985: 302) gives the following example as an unacceptable one: (i) *Mary is the person [whose teacher to consult (t) about that].}

(42) a. %Rudy is a good person [whose brain to pick (t)].
    b. %an author [whose book to buy (t)]
    c. %The nominee [whose candidacy to support (t)] is Smith.

The relative pronoun whose can be analyzed as who+'s. The problem is the grammatical status of this possessive marker 's. Chomsky (1995: 263) claims that the possessive 's is the head D of DP, as shown in (43) below: 

\begin{align*}
\text{overt syntax.}
\end{align*}

\footnote{\text{Emonds (1985: 302) gives the following example as an unacceptable one: (i) *Mary is the person [whose teacher to consult (t) about that].}}
He argues that the impossibility of ‘left branch’ extraction of *whose* (leaving *book* behind) follows from the fact that *who* and *’s* do not form a constituent in (43).

This analysis, however, seems dubious. For one thing, such an analysis does not offer an account for other cases of left branch extraction violations:

(44) a. *Which, did John buy [NP_t book]?*  
    b. *How, is John [AP_t tall]?

The impossibility of left branch extraction in (44) clearly does not show that *which* and *how* are not constituents. We will need some form of ‘Left Branch Condition’ to rule out (44). Once we assume this, the impossibility of left branch extraction of *whose* will presumably follow from this condition, and will no longer constitute a piece of evidence that *whose* is not a constituent.

There is also a piece of direct evidence against the analysis in (43). The following sentences are acceptable:

(45) a. I will read either [Chomsky’s] or [Lasnik’s] book.  
    b. I do not know [either [whose] or [whose wife’s] lawyer],  
      (t) undertook the divorce case.

The possibility of coordination in (45) above shows that NP+*’s* such as *whose* is a constituent, contrary to Chomsky’s claim.32

32 Chris Tancredi (personal communication) suggested that (45a) might be derived by ‘across the board rightward movement’ (or Right Node Raising) of *book*, as shown in (i):

(i) either [Chomsky’s (t)] or [Lasnik’s (t)] [book]

However, such an operation usually becomes possible only when the ‘moved’ element is heavy, as indicated in (ii):

(ii) I read [the first chapter of Chomsky’s (t)], not [the second chapter of Lasnik’s (t)], {??[book]  
    {book about syntactic structures}.

---

(43) \[
\begin{array}{c}
\text{DP} \\
\text{who} \\
\text{D'} \\
\text{D} \\
\text{N(P)} \\
\text{’s} \\
\text{book}
\end{array}
\]
What, then, is the categorial status of this possessive marker 's? Compare the following pairs of sentences:

(46) a. John knows a man \[ \text{whose father} \]

    \[ \text{the father of whom} \]

    is a lawyer.

b. John is looking for a car \[ \text{whose roof} \]

    \[ \text{the roof of which} \]

    is red.

Clearly, the possessive 's plays the same role as the preposition (of) in (46) above. We claim that the possessive 's is a kind of postposition, which is exceptional in English. This possibility is also suggested in Abney (1987).33

We have argued that the null head P of an IR attracts the \([+P]\) feature covertly. A possible derivation of (42a) is given in (42') below:

(42')

\[
\begin{array}{c}
\text{PP} \\
\text{CP} \\
\text{P} \\
\phi
\end{array}
\]

\[
\begin{array}{c}
\text{NP} \\
\text{N} \\
\text{PosP} \\
\text{Pos brain} \\
\text{who} \\
\text{'s} \\
\text{[+Adp]}
\end{array}
\]

\[
\begin{array}{c}
\text{C} \\
\text{IP} \\
\text{NP} \\
\text{(PRO)} \\
\text{I'} \\
\text{to}
\end{array}
\]

An anonymous reviewer has commented that (45b) is 'rather awkward.' According to Chris Tancredi (personal communication), (45b) is syntactically fine but 'pragmatically awkward,' in that it is difficult to think of a context in which it would be a plausible thing to say.

33 Strictly speaking, Abney (1987: 81–85) analyzes the possessive 's as a "postpositional case-marker," and distinguishes it from "true" adpositions. Even if Chomsky's analysis of 's as a D were to be retained, it might still be maintained that this D has a special property of having a feature \([+adposition]\), and a similar account of (42) would be possible.

34 The label Pos(P) in (42') can be taken to stand for either 'Possessive (Phrase)' or 'Postposition (Phrase). We use the term 'Possessive' for convenience, even though semantically 's marks a much wider range of relations than simple 'possession.'
For speakers who accept IRs like (42), the null head P of an IR (covertly) attracts the feature [+Adp(osition)], which includes both [+Preposition] and [+Postposition]. The feature [+Postposition] of the possessive 's is attracted to the null head P above, and the requirement in (20b) is satisfied by this feature. For speakers who do not accept IRs like (42), on the other hand, the null head P of an IR attracts only the feature [+Preposition]. Since the possessive 's has the feature [+Postposition], it cannot be attracted to the null head P position. The only candidate for satisfying (20b) is thus the feature [+Preposition] of the infinitival to. In order for the feature of to to be able to move to the null head P position, however, to must first move to the intervening head C position, or else the movement will violate the HMC.\(^{35}\) Once the infinitival to moves to the head C position, a 'doubly filled comp' violation results, and the structure is ruled out by (28'), as were (2a'') and (35b) above.\(^{36}\)

Thus, the idiolectal variation with respect to the acceptability of IRs with possessive relative pronouns like (42) can be reduced to an idiolectal difference in the lexical property of the null head P which forms IRs. There is parametric variation among speakers with respect to the feature that the null head P of an IR attracts: [+Adposition], or [+Preposition].

5. Concluding Remarks

5.1. Conclusion

We have shown that IRs in English can be analyzed as PPs headed by a null P. We argued that the null head P has a 'weak' feature which attracts the feature [+P] covertly. We claimed that the null affix C in IRs can induce overt syntactic head movement of the infinitival to to the C position only when the movement is required for convergence. We have shown that certain peculiar properties of IRs and the behavior of various types IRs which have defied satisfactory explanations follow straightforwardly under our analysis from the interac-

\(^{35}\) We are assuming here that 'covert head movement' of a feature also conforms to the HMC: the feature(s) of C might act as an 'intervening head' for covert movement of the feature [+P]. See also footnote 16.

\(^{36}\) Sentence (i) below is ruled out by (28') for any speaker.

( i ) *Rudy is a good person [whose brain for [John to pick (t)]].
tion of the Head Movement Constraint, and the ‘Doubly Filled Comp’ effect as a condition on spec-head agreement.

5.2. Remaining Problems

IRs have some other peculiar properties that we have not discussed in this paper. Berman (1974a, b) points out that the acceptability of an IR (within an object NP) is affected by the choice of the main verb:

(47) *John lost a book [for Mary to read (t)]. (cf. (1b))

We have not dealt with IRs in which the ‘gap’ corresponding to the head noun is in the subject position:

(48) John found [a man [(ec?) to do the job]].

According to Chomsky and Lasnik (1977), such sentences become less acceptable when this type of IR appears within an NP in the subject position:

(49) ?[A man [to fix the sink]] is at the front door. (cf. (14))

Ross (1986:231) points out that the sentence becomes degraded when the trace in an IR is embedded, especially in a tensed clause:

(50) Here is a knife [for you to (?say that you) cut up the onions with (t)].

In this paper we have not offered sufficient explanations for these properties of IRs, and would like to leave them for future research.

37 Williams (1980) and Browning (1987) (see also Hasegawa (to appear) for a similar claim) argue that the ‘empty category’ in IRs with a subject ‘gap’ is a PRO, rather than a trace of an empty operator. The absence of operator movement from the subject position of an IR might be related to the absence of ‘Tough-movement’ from the subject position, which might follow from the fact that variables must appear in (‘non-null’) Case positions:

(i) *John is tough [Op t [i to like Mary].
See also Bošković (1997: 186) for a possible analysis of IRs with a PRO as a ‘subject gap,’ which he attributes to Andrew Radford.

38 The degradedness of embedding in an IR might be related to a ‘mismatch of features’ caused by the movement of a relative operator in an IR to (the Spec of) an intermediate Comp, under the assumption that the movement of a relative operator in an IR is successive cyclic. We need to clarify what exactly the relevant features are of the relative operator in an IR and of the intermediate head C, and why such ‘mismatch of features’ results in lower acceptability rather than in total ungrammaticality if mismatch of features is supposed to ‘cancel the derivation’ (cf. (29)). We leave these problems open here.
REFERENCES


Fukui, Naoki (1995) “Comparative Syntax in English and Japanese: A Principles-and-Parameters Approach to Typology,” Approaches to Lan-


Hasegawa, Kinsuke (to appear) Bun I (Sentence I), Kenkyusha, Tokyo.


McCawley, James (1988b) “Adverbial NPs: Bare or Clad in See-Through Garb?” Language 64, 583–590.

McCawley, James (1991) Linguistic Flea Circus, Indiana University Linguistics Club, Bloomington, IN.


Ross, John (1986) Infinite Syntax! Ablex, Norwood, NJ.


Department of English
Tokyo Metropolitan University
Minami Osawa, Hachioji-shi
Tokyo 192–0397
e-mail: hasegawa-hiroshi@c.metro-u.ac.jp