Non-successive-cyclic Movement

Masanori Nakamura

1. Introduction

In current minimalist syntax (Chomsky 2000, 2001, 2005), the cyclic nature of derivation is characterized as its interactions with the interfaces through the operation Spell-Out (also called Transfer) (see Uriagereka 1999). Spell-Out is triggered each time a phase is constructed. For Chomsky, who adopts the Phase Impenetrability Condition, cyclicity is a matter of syntax, as it had always been till the turn of the century. Departing from Chomsky’s conception of cycles, researchers such as Stjepanović and Takahashi (2001), Fox and Pesetsky (2005), Bošković (2007a) argue that cyclicity is essentially a PF-related phenomenon. In a nutshell, for them, moving elements must raise to the edge of a phase, because otherwise they would be trapped inside the spell-out domain and linearized in the wrong place, leading to a crash at PF.

Notice that the theory of cyclic linearization, whereby successive-cyclic movement is motivated by phonological considerations, makes a general, potentially far-reaching prediction. Specifically, we predict that elements with phonological content and those without it should behave differently with respect to cyclicity, simply because the former are subject to linearization, whereas the latter are not. It is from this theoretical perspective that the following remarks are made.² As a case study, we take up null operators here. Lacking phonological features, they are expected to behave differently from phonologically overt elements like wh-phrases. It is
shown that their puzzling syntactic characteristics can be accounted for by an analysis based on Bošković's (2007a) treatment of Move and (long-distance) Agree. The success of the analysis in turn lends empirical support to the theory of cyclic linearization.

The rest of this paper is structured as follows. Section 2 considers the issues of locality imposed on Move and Agree. In particular, we examine how these two operations differ in terms of successive cyclicity, minimality, and islands, providing the basis upon which we can proceed to further discussion. Section 3 presents the predictions that the theory of cyclic linearization makes about the syntactic behavior of null operators. Section 4 discusses relevant data on null operator movement. It turns out that null operators have mixed properties of Move and Agree with respect to locality effects. Such properties prompt us to adopt an analysis, whereby null operators move in one fell swoop to their licensing positions, which we suggest are their probe heads. Finally, section 5 touches upon some implications of the present analysis and concludes the paper.

2. Locality of Move and Agree

Since Chomsky 1973, 1977, it has been a common wisdom that movement operates in successive-cyclic fashion. Thus in (1),

(1) Who does John think \([_{CP} t']_i [that Mary met t_i]]\)?

the \(wh\)-element who passes through the embedded Spec of CP on its way to the sentence-initial position. One particularly straightforward argument for the kind of derivation depicted in (1) comes from such examples as (3).

(2) *John, thinks that Mary likes pictures of himself,.
(3) [Which picture of himself, does John think \([_{CP} t']_i [that Mary likes t_i the best]]\)?

(2) is excluded for a familiar reason: it violates Principle A of Binding
Non-successive-cyclic Movement

Theory, which demands reflexives like *himself* be bound by their antecedents in local domains (in the case of (2), the embedded clause). On the other hand, (3) with the *wh*-movement of the DP containing *himself* is grammatical. It shows that the movement “rescues” the otherwise ill-formed sentence. To be more specific, the reflexive can be bound by the matrix subject *John* from the intermediate landing site higher than the embedded subject *Mary* (see Chomsky 1995 for an analysis). The proliferation of binding possibilities of the sort shown in (3) indicates that *wh*-movement indeed proceeds successive cyclically.⁴

Chomsky (1973) proposed that the successive-cyclic nature of movement is attributable to the Subjacency Condition, a syntactic condition which prohibits things from “moving too far.”⁵ The impact of this influential proposal has lasted for a long time and successive cyclicity has commonly been taken to reflect a unique property of the syntactic component (in this regard, Chomsky 2001 is no exception; see below).

However, recent years have witnessed the emergence of works (Stjepanović and Takahashi 2001, Fox and Pesetsky 2005, Bošković 2007a among others) arguing that successive cyclicity has its roots in the syntax-phonology interface, more specifically, in the way movement interacts with multiple Spell-Out (Uriagereka 1999). Consider the following configuration where H is a phase head:

(4)  \[[H \_P \_X \_P \_ [H \_Y \_P]]\]

According to Chomsky (2000, 2001), phase heads are *v* and *C*. The Phase Impenetrability Condition (PIC) (Chomsky 2000, 2001) has the effect of making the complement of the head (YP) inaccessible for further syntactic computations: only the head and its edge (XP) remain syntactically active.⁶ The above-mentioned works claim that the PIC can be dispensed with, given that linearization of syntactic structure takes place cyclically in accordance with multiple Spell-Out: once *YP* in (4) is sent to the phonological component and the terminal elements contained in *YP* get linearized, none of them would be able to move out of *YP*, because
typically that would lead to contradictory ordering statements in PF (see Fox and Pesetsky 2005, Bošković 2007a for details). It follows that successive-cyclic movement, forced by considerations of linearization, must be through the Specs of phase heads. Then there is no need to posit the PIC as an independent syntactic locality condition.7

If phases are truly irrelevant for syntactic locality, we would expect that they do not constrain syntactic operations such as Agree. Stjepanović and Takahashi 2001 and Bošković 2007a argue that that is precisely the case, citing examples such as (5) from Chukchee, a Chukokto-Kamchatkan language spoken in northeastern Siberia (Inènlıkèj and Nedjalkov 1973 and Mel’čuk 1988).8

(5) enan qêlyižu lerørke-nin-et [CP iŋqun pro Ø-retemŋ̃ev-nen-at
he-INST regrets-3-PL COMP 3SG-lost-3-PL
qora-t].
reindeer-PL(ABS)
‘He regrets that he lost the reindeers.’

In (5) the matrix verb agrees in number with the absolutive nominal in the embedded clause. Notice the presence of the complementizer iŋqun, showing that the agreement has taken place long-distance across the CP phase. The comparison between English (1) and Chukchee (5) leads to the following observation:

(6) Move exhibits PIC effects, whereas Agree does not.

The validity of (6) implies that the account of successive cyclicity based on cyclic linearization is superior to the one based on the PIC as a syntactic condition.

Before proceeding further, let us be explicit about when we need to invoke Move and/or Agree. Working in a derivational framework where no look-ahead beyond a phase is allowed, Bošković 2005 argues that Move
Non-successive-cyclic Movement

applies to an element with an uninterpretable feature (cf. the Activation Condition proposed in Chomsky 2000) as a last resort. He adopts a version of Last Resort, where X can undergo movement if and only if without the movement, the structure would crash. He also assumes with Epstein and Seely 2006 that the goal with an uninterpretable feature must function as a probe for its c-commanding probe with a matching uninterpretable feature and that an uninterpretable feature of X can be checked and deleted if and only if X c-commands the checker. It follows then that an element with an uninterpretable feature must move to the edge of a phase if the feature cannot be checked and deleted within the phase. Given the Shortest Move requirement, the final destination for the moving element will be the closest position c-commanding the probe. Note that the final step of successive-cyclic movement differs from the preceding steps in that it involves not only Move but also Agree: it achieves mutual feature checking between the probe and the goal through Agree.

Agree alone, on the other hand, is invoked when an element with an uninterpretable feature c-commands another fully interpretable element it agrees with. A typical case is (5) from Chukchee, where the matrix verb ‘regret’ with the uninterpretable agreement feature c-commands and agrees with the absolutive noun phrase ‘reindeers’ with the interpretable number feature.

Let us now consider the locality of Move and Agree in terms of minimality (cf. Rizzi 1990). It is well known that Agree is subject to the closest c-command requirement (Chomsky 2000, 2001). Consider (7) from Tsez, a Nakh-Dakhestanian language spoken in the northeast Caucasus (Polinsky and Potsdam 2001):

(7) eni-r [uz-a magalu b-ac-ru-li] b-i-y-xo.
 mother-DAT [boy-ERG bread.Ⅲ.ABS Ⅲ-eat-PSTPRT.NMLZ Ⅲ-know.PRES
 ‘The mother knows that the boy ate the bread.’

In (7) the matrix verb agrees long-distance in noun-class (i.e., class Ⅲ) with the absolutive nominal in the embedded clause. Putting aside the
question of word order (note that scrambling is rather free in Tsez), let us assume following works such as Murasugi 1992, Nakamura 1996, Mclachlan and Nakamura 1997 among others that the absolutive should be equated with the nominative and is the highest nominal in a given clause. Then it is expected that in examples like (7), only the absolutive nominal can agree with the matrix verb. This expectation is fulfilled: in Tsez non-absolutive DPs in the embedded clause never trigger agreement on the matrix verb, as shown, for example, in (8):

    mother-DAT [boy-ERG girl.II-DAT bread.ABS give-PSTPRF.NMLZ] II know
    (‘The mother knows that the boy gave the girl bread.’)

In the above example, the matrix verb agrees in class with the dative. The Agree Closest requirement prohibits the matrix verb from agreeing with the non-absolutive DP and hence (8) is ruled out.

Turning to Move, if it is correct to claim that Move, unlike the syntactic operation Agree, is motivated essentially for phonological reasons, we should expect Move to behave differently from Agree with respect to minimality effects. In particular, Move is expected not to exhibit such effects. In this connection, observe the following French examples, discussed by Bošković (1988):

(9) a. Tu as vu qui?
    you have seen who
    ‘Who did you see?’

   b. Qui as-tu vu?
      who have-you seen
      ‘Who did you see?’

(10) a. *Jean et Pierre croient que Marie a vu qui?
    Jean and Pierre believe that Marie has seen who
    (‘Who do Jean and Pierre believe that Marie saw?’)
b. Qui Jean et Pierre croient-ils que Marie a vu?
who Jean and Pierre believe-they that Marie has seen
‘Who do Jean and Pierre believe that Marie saw?’

(9a) shows that French in principle permits matrix \(wh\)-in-situ. Its variant with \(wh\)-movement ((10)) is also fine. Interestingly, a sharp contrast between \(wh\)-in-situ and \(wh\)-movement emerges when we consider embedded \(wh\)-elements. In particular, long-distance \(wh\)-in-situ, as in (10a), is disallowed, whereas long-distance \(wh\)-movement, as in (10b), poses no problem. Bošković (1988) analyzes \(wh\)-in-situ in French as involving Move F (Chomsky 1995). In his more recent work (Bošković 2007a), he proposes to reanalyze it as involving Agree instead. The main thrust of his proposal is that (10a) is a case of intervention in the necessary Agree relation between the matrix C and the \(wh\)-in-situ: the embedded C (phonetically realized as \(que\)), lying between the probe and the goal, blocks long-distance agreement. Crucial for this proposal is the idea that the exact specification of the feature involved is not important. Thus, the intermediate C with the specification [-wh] has the potential to block \(wh\)-feature checking. (9a) is well-formed because no C intervenes between the \(wh\)-in-situ and its probe. This analysis also accounts for the grammaticality of (10b). In (10b) Move raises the \(wh\)-phrase to a position higher than the embedded C, allowing the matrix C to establish an Agree relation with the \(wh\)-phrase. These considerations lead to the following thesis:

(11) Agree exhibits minimality effects, whereas Move does not.

(11), like (6), is only natural, given that Move, driven by phonological requirements, and Agree, utilized exclusively in syntax, are fundamentally distinct operations.

Shifting our attention now to island effects, Move is known to exhibit them, as in (12) (see Ross 1967).

(12) *How many cities does Mary have sisters who live in it?
In (12) *how many* has been extracted in violation of the Complex NP Constraint, and the result is ungrammatical.

What about Agree? Here again, we would expect it to differ from Move, given that they have nothing in common with respect to their motivations: it would be redundant and thus theoretically undesirable if the two operations are constrained by the same set of rules. But empirically, it seems difficult to find relevant data, because Agree is, in a sense, more local than Move: as we observed above, it is subject to minimality effects, and islands are likely to contain some interveners for the purpose of Agree.

Besides the redundancy argument, we have good reason to think that islands are irrelevant to Agree. To see the nature of island violations, let us consider Merchant’s (in press) analysis of island repair (Lasnik 2001, Fox and Lasnik 2003). Observe the example of sluicing in (13), taken from Lasnik 2001 (phonological deletion is indicated by strikethrough).

(13) Every linguist, met a philosopher who criticized some of his work, but I’m not sure [\_CP how much of his work \_TP every linguist, met a philosopher who criticized \_] .

In (13) the *wh*-phrase contains the variable *his*, bound by the universal quantifier *every linguist*. The presence of the bound variable, which must reconstruct into a position c-commanded by its binder at LF, ensures that the sluicing site contains the complex NP island. Thus the grammaticality of (13) shows that the PF-deletion saves the otherwise ill-formed structure with an island violation (compare (12) and (13)).

Chomsky (1972) explains island repair by assuming that crossed island nodes are marked with a PF-uninterpretable feature. His account correctly predicts that (13) is well-formed: since the entire island is deleted, no PF-defective node is left behind. This account, however, faces empirical problems, as pointed out by Merchant (in press). Consider (14).

(14) a. Does Abby speak the same Balkan language that *Ben* speaks?
b. *No, \([_{FP}Charlie]_{CP} t_{TP}\) Abby speaks the same Balkan language that \(t\) speaks].

(14b) is an example of Hankamer’s (1979) “wrong” transformation, where the second speaker makes a correction to some aspect of the first speaker’s utterance. In (14) what is corrected is *Ben within the complex NP island. Merchant (2004) argues that in (14b) the focused nominal *Charlie moves to what he agnostically calls Spec FP above CP and then the sentence undergoes TP-ellipsis. Contrary to the prediction made by Chomsky’s (1972) analysis, eliminating the entire island does not save the structure. Merchant (in press) presents an alternative account, in which it is traces/copies of island-escaping elements rather than island nodes that are marked with a PF-uninterpretable feature. Under this account, the TP-ellipsis necessarily leaves behind the CP projection containing a defective trace/copy (indicated by \(*t\)), leading to a PF-crash. (13), on the other hand, is grammatical, since the sluicing gets rid of all the PF-defective traces/copies. Therefore, Merchant’s account is empirically superior to Chomsky’s.

Note that the success of Merchant’s analysis implies that only Move is subject to island effects, because island violations arise only when we have PF-illegitimate traces/copies. It must be then that Agree, which leaves no traces/copies, is immune to island effects. We thus conclude that the following holds:

(15) Move exhibits island effects, whereas Agree does not.

To summarize this section, Move and Agree are constrained by different locality conditions. The former exhibits PIC and island effects, and the latter minimality effects. This state of affairs is only natural, given that the two operations are essentially driven by distinct needs: Move is designed to meet phonological requirements, whereas Agree is used to eliminate uninterpretable features in syntax.
3. Predictions

With the preceding discussion of the locality of Move and Agree in mind, let us consider what kinds of predictions the theory of cyclic linearization makes with respect to null operators, which are of interest for the following reason. As is well known, null operators participate in movement-like dependencies (see Browning 1987 among others). In addition, from the viewpoint of cyclic linearization, they are expected to behave differently from regular wh-operators, precisely because they lack phonological content and thus do not have to be linearized in the first place. We can think of two possible analyses, both of which are mentioned in Takahashi 2001.⁹

First, one may think that since null operators are free from linearization, they are not subject to Move at all. Under the standard GB analysis, the “tough-sentence” in (16) has the representation in (17), where $OP$ stands for null operator:

(16) John is easy to convince Bill to work with.

(17) John is easy $[_{CP} OP_{i} [_{TP} PRO$ to convince Bill $[_{CP} t'_{i} [_{TP} PRO$ to work with $t_{i}]]]]$

In (17) the null operator, just like a wh-phase, undergoes successive-cyclic movement via the intermediate Spec of CP. In contrast, under the non-movement analysis, (16) has the following representation:

(18) John is easy $[_{CP} C [_{TP} PRO$ to convince Bill $[_{CP} C [_{TP} PRO$ to work with $OP]]]]$

In (18) the null operator stays in situ and gets associated with the highest C in the complement of easy by Agree (recall that Agree can apply long-distance across phase boundaries).¹⁰ Note that within the framework of Bošković 2005, this analysis presupposes that the null operator does not have an uninterpretable feature: only the C does, and that is why Agree suffices in (18).
Second, one may choose to maintain the idea that null operators do have an uninterpretable feature. Then they would be forced to undergo movement to a position from which they c-command their probe, as discussed above. Even under this assumption, the theory of cyclic linearization would characterize the syntactic behavior of null operators in an interesting way. In particular, it would predict that null operator movement is in one fell swoop: there is no need for a null operator lacking phonological content to move to the edge of an intermediate phase for the purpose of linearization. The one-fell-swoop analysis is illustrated in (19).

(19) John is easy \([_{CP} OP_{i} [_{TP} PRO \text{ to convince } Bill [_{CP} C [_{TP} PRO \text{ to work with } t_i]]]]\)

Unlike (17), (19) does not contain an intermediate trace/copy and the null operator moves directly to the vicinity of C heading the entire embedded clause.\(^{11}\)

Given that the theory of cyclic linearization is superior to the phase theory based on the PIC, the question now is: which representation is correct, (18) or (19)? Notice that the non-movement analysis and the one-fell-swoop analysis can be teased apart empirically. The former involves Agree only, whereas the latter uses a combination of Move and Agree (by definition, Move entails Agree, since a moving element has an uninterpretable feature to be checked and deleted by Agree).

4. Null Operator Movement

Let us now examine the syntactic behavior of null operators in order to choose between the non-movement analysis and the one-fell-swoop analysis. Making the choice, it turns out, is not difficult. Consider the following examples ((21) is adapted from Browning 1987:22):

(20) *John is easy to believe [(that) Mary kissed \( t \)].
(21) *John is easy to describe to Bill a plan to assassinate \( t \).
As shown in (20), one characteristic of “tough-movement” is that it cannot take place out of a tensed clause (Stowell 1986 among others). In ill-formed (21), the null operator is extracted out of a complex NP headed by a plan. Note that the reason for the ban on the extraction must be different from the one for the ill-formedness of (20): there is no tensed clause in the complement of easy in (21). (21) then shows that null operator movement obeys island constraints.

The fact that null operator movement is island-sensitive counts as evidence for the superiority of the one-fell-swoop movement analysis, because as mentioned above, island sensitivity is a trait of Move, not Agree. Thus (20) and (21) have the representations in (22) and (23) respectively.

\[(22) \ast \text{John is easy} \ [_{\text{CP}}OP_{i} \ [_{\text{TP}} \text{PRO to believe} \ [_{\text{CP}} \text{(that)} \ [_{\text{TP}} \text{Mary kissed} \ t_{i}]]]]\]
\[(23) \ast \text{John is easy} \ [_{\text{CP}}OP_{i} \ [_{\text{TP}} \text{PRO to describe to Bill} \ [\text{a plan to assassinate} \ t_{i}]]]]\]

(23) is ruled out since the trace/copy in the object position of assassinate, left by the island-violating movement, has a PF-uninterpretable feature. What is wrong with (22)? There is no problem with Move here, because the null operator movement does not have to proceed successively cyclically and does not cross an island, either. Then there must be a problem with Agree. Let us assume that in (22) the probe C has an uninterpretable feature, call it op-feature. Lasnik and Stowell (1991) argue convincingly that unlike wh-phrases, null operators of the kind we have been considering are not quantificational and their traces/copies are null R-expressions with binding properties of names and definite descriptions (“epithets”). Thus they are of the category DP, in need of identifying their antecedents. Assume further on the basis of the similarities between D and T (Enç 1987 and Hornstein 1990) that finite T (but not nonfinite one) also has a species of the feature in question (Nakamura 2002a). Given these assumptions, it follows that the embedded tensed T, no matter what the exact specification of its op-feature is, blocks long-distance agreement between the probe C and its goal, represented by \(t_{i}\) in (22). This leads to a crashed derivation, with the goal’s
uninterpretable *op*-feature unchecked.\textsuperscript{12}

The present analysis also explains the (relative) well-formedness of (24) (taken from Authier 1989), which has the derivation in (25).

(24) ?John is easy to know what present to give to.
(25) ?John is easy $[_{\text{CP}} OP, [_{\text{TP}} \text{PRO to know} [_{\text{CP}} \text{what present}_j [_{\text{TP}} \text{PRO to give } t_j \text{ to } t_i]]]]$

As shown above, a null operator can in principle be extracted out of a *wh*-island. This may be perplexing, given that null operator movement is sensitive to other kinds of islands, as shown in (21). The key to understanding the grammaticality of (24) lies in the nature of *wh*-islands: *wh*-island effects are in fact PIC effects, whereas other island effects are arguably not (they fall under Condition on Extraction Domain (Huang 1982) effects). Consider the typical *wh*-island violation in (26) (Authier 1989:122).

(26) *Which car did she know $[_{\text{CP}} \text{what}_j [_{\text{TP}} \text{PRO to put } t_j \text{ on } t_i]]$?

The extraction of *which car* out of the embedded clause is prohibited because as soon as *what* reaches the embedded Spec of CP, the embedded TP is sent to PF for linearization and *which car* is trapped inside the TP.\textsuperscript{13} *Which car* has no chance of moving to the edge of the embedded CP since English C does not license multiple specifiers.

In contrast to *wh*-movement, null operator movement does not exhibit PIC effects. This is why in (25) the long-distance movement of the null operator is legitimate. In addition, no finite T lies between the trace/copy of the null operator and its probe C, allowing Agree to establish the necessary association between the two elements. In short, there is no violation of locality conditions in (24).

The one-fell-swoop analysis predicts that null operators can participate in long-distance dependencies if nothing induces minimality effects. It correctly predicts the grammaticality of examples like (19), where
no finite T intervenes between the probe C and the trace/copy of the null operator. In addition, it comes to grips with the syntactic behavior of certain kinds of null operators that tolerate long-distance dependencies even across finite T. Potts (2002b) argues that "as-parentheticals" of the kind exemplified in (27) involve null operator movement (Potts 2002b:637).

(27) Ames was a spy, as the FBI eventually discovered.

According to Potts's (2002b) analysis, the as-clause in (27) is derived as follows:

(28) \[
[_{pp} \text{ as } [_{CP} OP \ [_{TP} \text{ the FBI eventually discovered } t_{CP} ]] ]
\]

In (28) a null operator of the category CP undergoes movement into the specifier of the CP selected by as, analyzed as a preposition. Interestingly, CP null operators can be extracted out of a finite clause, as shown in (29) (the relevant portion of its derivation is illustrated in (30)) (Potts 2002a,b).

(29) We should resign right away, as I'm sure you'll agree.
(30) \[
[_{pp} \text{ as } [_{CP} OP \ [_{TP} \text{ I'm sure } [_{CP} C \ [_{TP} \text{ you'll agree } t_{CP} ] ] ] ]
\]

If we adopt the present analysis, we can begin to understand why DP null operators and CP ones behave differently in terms of finiteness. The question here is: why is it that in (30) (and in fact (28)), the finite T does not block long-distance agreement between the probe (which we assume is C) and the goal? The reason, we suggest, is simply that T lacks the relevant feature that is involved in the agreement: it is natural to assume that the relevant feature (which may be called "propositional op-feature") has nothing to do with the notion of specificity/definiteness (Enç 1987), associated with D and T. It follows automatically, then, that only DP null operators are subject to minimality effects exerted by finite T. The contrast between (20) and (29) counts as evidence for the superiority of the present feature-based account making use of Agree over other approaches.14

As predicted, since as-parentheticals involve null operator movement,
they are sensitive to islands. An example of an island violation is given in (31) (Potts 2002b:631).

(31) *Americans have a right to cheap gas, just as George espoused his belief that the world should accept.

The derivation of the *as-clause in (31) is shown below:

(32) *[P as [CP OP₀ [TP George espoused [his belief [CP that [TP the world should accept tCP]]]]]]

(32) is excluded on a par with (23), because the trace/copy, represented by tCP, is marked with a PF-uninterpretable feature.¹⁵

Above we saw an example where null operator movement is not constrained by wh-islands (see (24)). In this light, the ungrammaticality of (33) may seem surprising (Potts 2002b:632).

(33) *Chuck rides a unicycle, just as Sue asked me whether I knew.

In (33) the null operator has been extracted out of the wh-island, as illustrated below:

(34) *[P as [CP OP [TP Sue asked me [CP whether [TP I knew tCP]]]]]

Note that the movement of the CP null operator is not affected by the presence of finite T, as demonstrated in (30). Then is the ill-formedness of (34) due to a wh-island violation? We suggest it is not, in light of the above theoretical as well as empirical considerations that point to the absence of PIC effects in the case of null operator movement. If there is nothing wrong with the movement in (34), the source of the ungrammaticality must be failed long-distance agreement. More specifically, it must be that the wh-C prevents the probe C from agreeing with the goal. As mentioned earlier, the relevant feature can be characterized as “a propositional op-feature” since the probe searches for an operator of the category CP. Suppose, as seems
reasonable, that the *wh*-operator feature, associated with C, and the propositional *op*-feature are of the same kind, although they differ in their exact specifications. The minimality effect observed in (34) is now expected under Bošković’s (2007a) theory of Agree. In belief, we argue that the variable sensitivity to *wh*-islands shown by different kinds of null operators provides additional support for the proposed analysis.\(^\text{16}\)

To recapitulate, we have seen that null operator movement exhibits hybrid behavior with respect to the locality effects Move and Agree are subject to, as summarized in (35).

\[
\text{(35)} \quad \begin{array}{cccc}
\text{PIC effects} & \text{Move} & \text{Agree} & \text{OP Mov’t} \\
\text{minimality effects} & \text{yes} & \text{no} & \text{no} \\
\text{island effects} & \text{no} & \text{yes} & \text{yes} \\
\end{array}
\]

Null operator movement is similar to Agree, but not Move, in that it exhibits minimality effects but does not show PIC effects. At the same time, it is similar to Move, but not Agree, in being island-sensitive. Its mixed behavior, we argued, stems from the lexical properties of a null operator, that is, it has an uninterpretable feature and, importantly, lacks phonological content. As a consequence, it undergoes one-fell-swoop movement. It avoids PIC effects, because it is phonologically null and thus free from cyclic linearization. It is subject to minimality effects, because it involves long-distance Agree: unlike elements with phonological features, it cannot move to a non-feature-checking position higher than a potential intervener. Furthermore, it is constrained by islands, because it does involve movement.

Having established that null operators do undergo (one-fell-swoop) movement, let us now consider what exactly their landing sites are. So far, we have assumed along the lines of Epstein and Seely 2006 that null operators move to Spec of CP. It has been argued, however, that for the purpose of feature checking, head adjunction is preferred over movement to a specifier position (Nunes 1998, Bošković 2001). Since null operators,
void of phonological content, cause neither a morphophonological problem nor a linearization problem with their probes, they are supposed to take the default option. Adopting the Move F framework of Chomsky 1995, Nakamura (2002a) argues that this is precisely the case based on the interactions between phrasal *wh*-movement and null operator movement.

We have seen in (26) above that a *wh*-phrase cannot be extracted out of an infinitival *wh*-island. This is because the *wh*-phrase in the embedded Spec of CP prevents the other one from moving to the edge of the embedded CP phase, leaving the latter trapped inside the embedded clause. The head adjunction analysis of null operator movement predicts that *wh*-movement can take place out of a CP whose head is a probe of a null operator: the *wh*-phrase in question should be able to move to the edge of the CP, which is not filled by a null operator. This prediction is in fact borne out. Compare (26) with the following example (Authier 1989:122 citing Kirkpatrick 1982):

(36) Which car, did she buy those whitewallsj[cP t′i OPj C [TP PRO to put tj on tj]]?

In (36) the null operator is assumed to adjoin to its probe for feature checking. The *wh*-phrase which car undergoes usual successive-cyclic movement via the available embedded Spec of CP. Data like (36) provide evidence for the claim that null operator movement is head adjunction.17

In short, based on the recent theory on the locality of Move and Agree, we have proposed that null operators move in one fell swoop to adjoin to their probes. This proposal accounts for the locality conditions imposed on null operator structures as well as the interactions between null operator movement and *wh*-movement.

5. Implications and Conclusion

To the extent that the preceding discussion of the behavior of null operators is on the right track, it provides additional support for the theory
of cyclic linearization advocated by Stjepanović and Takahashi 2001, Fox and Pesetsky 2005, Bošković 2007a and others. It implies that as argued by these authors, the Phase Impenetrability Condition as a syntactic condition should be dispensed with. The locality of Move is imposed by the mechanism of cyclic linearization, couched within the theory of phase, whereas Agree, a purely narrow-syntactic operation, is constrained by the minimality condition (Agree Closest).

The theory of cyclic linearization contends that no feature checking takes place in intermediate positions that successive-cyclic movement goes through (Bošković 2005, 2007b). In view of the present analysis of null operator movement, the following example from Irish offers direct evidence for this contention (McCloskey 2004).

(37) Bhí lá galánta ann mar a thuar Proinnsíos a bheadh t
    was day beautiful in-it as aL predicted Proinnsíos aL be(COND)
    'It was a beautiful day, as Proinnsíos had predicted it would be.'

(37) contains the Irish equivalent of the English parenthetical _as_-construction discussed by Potts (2002a,b). Notice the presence of the two instances of the so-called _wh_-agreement morpheme _aL_, which has traditionally been taken to provide one of the strongest pieces of evidence for the successive-cyclic nature of _wh_-movement and the feature checking (the checking of the generalized EPP feature, in particular) in intermediate landing sites (see McCloskey 1990 among others). The present analysis gives the _mar_-clause in (37) the following schematic representation:

(38) [pp mar [cp OP-a [thuar Proinnsíos [cp a [bheadh t]]]]]

Following Potts (2002a,b) and McCloskey (2004), we assume that a null operator is involved in the derivation of the _mar_-clause and that it must establish a checking relation with the complementizer of the clause selected by the preposition _mar_. In (38) the null operator moves in one fell swoop and adjoins to _a_, which we tentatively assume is a complementizer. (38), if
correct, demonstrates that the *wh*-agreement morpheme indicates neither Spec-head agreement nor feature checking. In fact, Noonan (2002) argues that the *wh*-agreement morpheme is not a special form of complementizer but a verbal morpheme signaling “object shift.”

The analysis of null operator movement defended here also argues for the thesis that derivational operations (especially, Move) are sensitive to the presence/absence of phonological features. This clearly runs counter to one of the leading ideas of the Distributed Morphology (Halle and Marantz 1993) that syntactic categories have no phonological content (Late Insertion). The present study corroborates the claim, supported by a growing amount of evidence for the deletion theory of ellipsis (see Lasnik 2001, Merchant 2001, 2004 among numerous others), that lexical items are introduced into derivations with phonological features (if they do have phonological features).

In conclusion, it has been argued that null operators move in one fell swoop to adjoin to their probe heads. They inherently have an uninterpretable feature and lack phonological features. These lexical properties, interacting with the workings of cyclic linearization of syntactic structure, are ultimately responsible for their peculiar behavior with respect to Move and Agree. This finding is fully consistent with the principles-and-parameters tradition, where it is maintained that there exists only one kind of computational system of human language and all superficial syntactic differences, intralinguistic or interlinguistic, are attributable to idiosyncrasies in the lexicon – a conceptually welcome result.

Notes

* This work was supported by a Senshu University Domestic Research Fellowship for the academic year 2007-2008, for which I am grateful.

1 Chomsky (2000:108) presents the following:
(i) Phase-Impenetrability Condition

In Phase $\alpha$ with head $H$, the domain of $H$ is not accessible to operations outside $\alpha$, only $H$ and its edge are accessible to such operations.

2 For discussion of phonologically overt vs. covert categories targeted by Spell-Out, see Nakamura in press.

3 Under the copy theory of movement, what is left by movement is a copy rather than a trace, but we use the trace notation throughout. In the framework of Chomsky 2000, 2001, wh-movement must proceed through the edge of phasal $vP$. We omit possible intermediate traces/copies in Spec of $vP$ in our examples, since as far as this paper is concerned, nothing important hinges on them.

4 For other pieces of evidence for successive cyclicity, see Fox 1999, McCloskey 2000 among many others.

5 Chomsky (1973) formulated the Subjacency Condition as in (i), where “bounding nodes” are $S'$ and $NP$:

(i) Subjacency Condition

No rule can involve $X$, $Y$ in the structure

...$X...[\alpha...[\beta...Y...]]...X...$

where $\alpha$, $\beta$ are bounding nodes.

6 See note 1.

7 The syntactic approach to successive cyclicity utilizing the PIC (Chomsky 2000, 2001) assumes that every movement to the phase edge is triggered by a (generalized) EPP feature. See Bošković 2005, 2007b for arguments that no feature checking is involved in intermediate positions that successive-cyclic movement passes through. See also Section 5.

8 The following abbreviations are used here:

<table>
<thead>
<tr>
<th>ABS-absolute</th>
<th>COMP-complementizer</th>
<th>COND-conditional</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT-dative</td>
<td>ERG-ergative</td>
<td>INST-instrumental</td>
</tr>
<tr>
<td>NMLZ-nominalizer</td>
<td>PRES-present</td>
<td>PSTPRT-past perfect</td>
</tr>
<tr>
<td>PL-plural</td>
<td>SG-singular</td>
<td>3-third person</td>
</tr>
<tr>
<td>III-class 3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Stjepanović and Takahashi (2001) and Bošković (2007a), from which (5) is taken, gloss the embedded object ‘reindeers’ as nominative, but in view of the fact that Chukchee is an ergative language, we gloss it as absolutive. For
arguments that both absolutive Case and nominative Case are licensed by T, see Murasugi 1992, Nakamura 1996, Mclachlan and Nakamura 1997 among others.

9 Takahashi’s (2001) discussion of null operators is built mainly around the hypothesis that they cannot satisfy the (generalized) EPP as a phonological requirement. Therefore, its relevance to the present paper is obvious.

10 One may well wonder why the intervening C does not block the association. It is suggested below that C, whether it is tensed or not, is not specified for the relevant feature involved in null operator licensing.

11 Later it is argued that null operator movement is head adjunction.

12 It has been pointed out that French-type wh-in-situ and null operator movement have similar distributional properties (see Nakamura 2002a for relevant discussion). For instance, they are both subject to the “Tensed-S Condition” (see (10a) and (20)). One may well wonder whether the present analysis of null operator movement extends to certain kinds of wh-in-situ. We leave a careful investigation of wh-in-situ from this perspective for future work. Note that even if the proposed analysis is shown to cover relevant wh-in-situ cases, the assumption that Agree is constrained by minimality remains valid.

13 The same kind of problem may arise at an earlier stage of derivation, at the embedded VP phase in particular (see note 3).

14 As expected, null operators of another non-DP kind are not subject to minimality effects at issue, either, as shown in (i) (Potts 2002b).

(i) He arrived on time, as I had said he would.

According to Potts (2002b), the as-parenthetical in (i) involves movement of a VP null operator (but see note 16).

It is worth pointing out the grammaticality of (29) in the text and (i) above provide evidence against the non movement analysis of null operators advocated by Takahashi (2001), who assumes that Agree is constrained by the Subjacency Condition.

15 What Potts (2002b) regards as a VP null operator is also sensitive to islands, just as expected (Potts 2002b:631):

(i) *Eddie fills his truck with leaded gas, just as they believed the report that he must.
Potts (2002b:632) mentions that (i) is excluded.

(i) *Chuck rides a unicycle, just as Sue asked me whether I could.

(i) shows that what Potts regards as a null operator of the category VP cannot move out of a wh-island. If the text analysis is on the right track, the ungrammaticality of (i), like that of (33), must be due to a minimality effect on Agree. We speculate that the null operator may be of the category vP rather than VP. In other words, it may count as a “propositional” category in the sense of Chomsky 2000, 2001 and have “a propositional op-feature” just like a CP null operator. If this is tenable, the account of (33) extends directly to (i).

One may think that the following example argues against the head adjunction analysis (Authier 1989:123):

(i) *Our military advisors should know \( [_{c_{\text{p}}} \text{ which missiles} ,_{t_{\text{p}}} \text{ these countries} , ] \) are too unfriendly \( [_{c_{\text{p}}} t' ,_{o_{\text{p}}} \text{ PRO to send} t , t ] \)\].

As pointed out by Nakamura (2002a), however, (i) does not count as a real counterexample for the analysis, because it is ruled out independently by some ban on crossing dependencies (see Pesetsky 1982).

See Nakamura 1996 for a similar view regarding Tagalog. Nakamura (in progress) attempts to analyze verbal wh-agreement as a way of establishing long-distance agreement between a null operator and its probe.

References


Chomsky Noam. 2005. On phases. Ms. MIT.


