【Article】

A Minimalist Approach to Non-Restrictive Relative Clauses in English*

TOZAWA Takahiro

1. Introduction

There are two types of relative clauses in English: restrictive relative clauses (RRCs) and non-restrictive relative clauses (NRCs). This is illustrated in (1a, b).

(1) a. He has two daughters who are studying music.
b. He has two daughters, who are studying music.

The clause introduced by *who* in (1a) is referred to as an RRC and it modifies its head *two daughters*. The modification of *two daughters* by the RRC restricts the referential content of daughters, limiting it to only the two daughters who are studying music. Thus, (1a) implies the possibility that he has another daughter who is not studying music. Conversely, the *who*-clause in (1b) is the NRC. This is a type of parenthetical expression that provides a supplementary explanation of its head. That is, the sentence in (1b) is paraphrased as "he has two daughters

* This study is an extended version of that of Tozawa (2020, 2022). I would like to express my appreciation of Yoshiaki Kaneko, Etsuro Shima, Toru Ishii, Satoshi Oku, Nobuhiro Miyoshi, Tadao Nomura, Yosuke Sato, Ryosuke Sato, and three anonymous reviewers for their invaluable advice and insightful comments. I would also like to thank Yoshihisa Goto, Satoru Kanno, and Masashi Totsuka for their inspiring discussions. I am also grateful to Christopher Bozek for providing me with English judgments. All errors are, of course, mine. This study was supported by JSPS KAKENHI Grant Number JP21K00564.

Studies in English Linguistics and Literature 33 (2023) 119-142
©2023 by the Japan Association of English Linguistics and Literature

and the two daughters are studying music." Therefore, (1b) implies that he has only two daughters. Thus, RRCs and NRCs differ in their functions. Further investigation of NRCs revealed that they have main- and subordinate- clause properties (reviewed in Section 3). This leads us to wonder how NRCs are syntactically derived. In this study, we propose that an NRC and main clause form a coordinate structure and that NRC antecedents undergo Sideward Merge (SM) or Parallel Merge (PM) with the NRC. Although Chomsky (2019a, b) and Chomsky, Gallego, and Ott (2019) argue that neither SM nor PM is legitimate, this argument holds only in the case of set-Merge by SM or PM. Whether pair-Merge by SM or PM is illegitimate remains to be addressed. We propose that SM and PM are possible if they are pair-Merge. In this proposal, we provide an SM- or PM- based analysis of NRCs.[1]

The remainder of this paper is organized as follows. Section 2 outlines the MERGE theory. Section 3 presents the previous analyses of NRCs. Section 4 proposes that SM and PM are possible if they are pair-Merge. Section 5 analyzes NRCs based on SM or PM. Section 6 discusses the consequence of the analysis. Finally, Section 7 concludes the study.

2. Theoretical Framework

The Minimalist Program is a research program aimed at building a linguistic theory based on the concept of the bare minimum. Under the research guidelines, Chomsky (2000) proposes the Strong Minimalist Thesis (SMT), as shown in (2).

(2) Language is an optimal solution to legibility conditions.
(Chomsky (2000: 96))

In line with the SMT, language is regarded as a perfect system in that it satisfies interface conditions in a way that accords with third-factor principles (e.g., the no-tampering condition). Assuming that Universal Grammar (UG) is a simple and elegant theory, Merge (which is the only UG operation) should be simple.

In the early Minimalist Program, Chomsky (1995) defines Merge as in (3).

(3) The simplest such operation takes a pair of syntactic objects (SO_i, SO_j) and replaces them by a new combined syntactic object SO_{ij}. Call this operation *Merge*. (Chomsky (1995: 226))

Merge applies to two syntactic objects, SO_i and SO_j, creating a new syntactic object SO_{ij}. In this way, Merge is an operation applied to two syntactic objects, and this is the simplest structure-building operation in the early minimalist framework. However, Chomsky (2019a,b, 2020) further pursues the simplest Merge, modifying its definition based on the discussion of subject-predicate structures. Specifically, the subject (NP) and predicate (VP) are formed independently before they are combined. This means that we need a place where the subject and predicate are built and combined. Chomsky (2019a, b, 2020) calls this place Workspace (WS) and redefines Merge as in (4).

(4) MERGE maps WS onto WS′. (Chomsky (2019a, b, 2020))

Merge applies to WS rather than syntactic objects and it maps WS into WS′. This version of Merge is referred to as MERGE. Chomsky (2019a) and Chomsky, Gallego, and Ott (2019) argue that MERGE has the property of Recursion. This is shown in (5).

(5) Recursion
All syntactic objects in the lexicon and in the workspace WS are *accessible* to MERGE; (...) The basic property of recursive generation requires that any object already generated be accessible to further operations. (Chomsky, Gallego, and Ott (2019: 245))

Recursion is a property of the Faculty of Language that requires all syntactic objects in WS to be accessible to MERGE. In addition, note that MERGE observes Minimal Search (MS), which is a third-factor principle. For example,

consider (6).

(6) $[_{CP}$ C ... what$_2$ ... what$_1$ [...]]

Suppose that *what$_1$* is a copy of *what$_2$*. If *what* undergoes movement to CP, *what$_2$* rather than *what$_1$* moves to CP. This is because MS first finds the higher copy *what$_2$* rather than the lower copy *what$_1$*, rendering *what$_1$* inaccessible to MERGE. MERGE, in addition to abiding by MS, is subject to Minimal Yield (MY) (Chomsky (2021)), as shown in (7).

(7) Minimal Yield (MY)
MERGE adds at most one new accessible term to an updated workspace.

When MERGE maps WS into WS′, the number of increased accessible terms must be at most one because of MY.

Thus far, we have seen that (i) MERGE changes WS into WS′, (ii) it has the property of Recursion, and (iii) it observes MS and MY. Let us consider the MERGE-based derivation of a sentence *John loves Mary*, as shown in (8).

(8) a. WS1 = [C, T, John, *v**, loves, Mary] (six accessible terms)
b. WS2 = [C, T, John, *v**, {loves, Mary}] (seven accessible terms)
c. WS3 = [C, T, John, {*v**, {loves, Mary}}] (eight accessible terms)
d. WS4 = [C, T, {John, {*v**, {loves, Mary}}}] (nine accessible terms)
e. WS5 = [C, {T, {John, {*v**, {loves, Mary}}}}] (ten accessible terms)
f. WS6 = [C, {John {T, {John, {*v**, {loves, Mary}}}}}]
(eleven accessible terms)
g. WS7 = [{C, {John {T, {John, {*v**, {loves, Mary}}}}}}]
(twelve accessible terms)

WS is a set of syntactic objects. Each WS in (8) contains lexical items and phrase structures.[2] WS1 contains six lexical items (C, T, John, *v**, loves, and Mary) and these syntactic objects are accessible to MREGE owing to Recur-

sion. MERGE is applied to WS1 to obtain WS2. WS2 contains a new syntactic object {loves, Mary} in addition to the six terms (C, T, John, *v**, loves, and Mary). Importantly, note that in the mapping from WS1 to WS2, the number of accessible terms increases by one. That is, it increases from six in WS1 to seven in WS2. This satisfies (7). Therefore, there is no problem with the derivation. Here, let me digress slightly to discuss why the lexical items *loves* and *Mary* in WS1 are eliminated in WS2. Consider (9).

(9) WS2′ = [C, T, John, *v**, loves, Mary, {loves, Mary}]

If MERGE gave rise to WS2′, it would violate (7). This is because the number of accessible terms increases by three. Note that WS2′ contains nine accessible terms: C, T, John, *v**, loves, Mary, loves of {loves, Mary}, Mary of {loves, Mary}, and {loves, Mary}. The number of accessible terms then increases from six to nine, violating (7). Therefore, in accordance with MY, *loves* and *Mary* are eliminated in the mapping from WS1 to WS2. In this way, we can eliminate *loves* and *Mary* from WS without recourse to *Remove*.[3] Let us now return to the derivation in (8). In (8c), MERGE transforms WS2 into WS3, forming a new syntactic object {*v**, {loves, Mary}} and eliminating *v** and {loves, Mary}. In WS4, the structure {John {*v**, {loves, Mary}}} is newly built and *John* and {*v** {loves, Mary}} are removed. In WS5, T is combined with {John {*v**, {loves, Mary}}} to form {T, {John, {*v**, {loves, Mary}}}}. In WS6, where *John* undergoes movement, accessible terms are eleven: C, John, T, *v**, loves, Mary, {loves, Mary}, {*v**, {loves, Mary}}, {John, {*v**, {loves, Mary}}}, {T, {John, {*v**, {loves, Mary}}}}, and {John, {T, {John, {*v**, {loves, Mary}}}}}. Note that only the higher copy of *John* is accessible because MS first reaches a higher copy. Therefore, the mapping from WS5 to WS6 satisfies (7). Finally, MERGE is applied to WS6 to obtain WS7. This is how the sentence is derived.

Chomsky (2019a, b, 2020) and Chomsky, Gallego, and Ott (2019) argue that under the theory of MERGE, PM (Citko (2005)), SM (Nunes (2001)), and Late Merge (Lebeaux (1991)) are illegitimate operations and should be excluded. Let us take up SM. Nunes (2001) proposes sideward movement, which is a

series of operations, as indicated in (10).

(10) a. $[_K \ldots \alpha_i \ldots]$ α_i ⟷ $[_L \ldots]$

Copy Merge

b. $[_K \ldots \alpha_i \ldots]$ $[_M \alpha_i [_L \ldots]]$ (cf. Nunes (2001: 305))

In (10a), α in K creates its copy and merges with the syntactic object L formed separately from K, creating the syntactic object M. This is SM.[4] However, SM violates (7). Consider the derivation in (11).

(11) a. WS1 = [a, b, c] (3 accessible terms)
b. WS2 = [{a, b}, c] (4 accessible terms)
c. WS3 = [{a, b}, {b, c}] (6 accessible terms)

We suppose that the derivation proceeds to WS2 in (11b), where *a* is merged with *b*. Here, *b* undergoes SM/PM with *c* to create {*b*, *c*} in WS3.[5] The number of syntactic objects in WS3 is six: *a*, *b* of {*a*, *b*}, *b* of {*b*, *c*}, *c*, {*a*, *b*}, {*b*, *c*}.[6] Given that all syntactic objects in WS3 are accessible by Recursion, SM/PM violates (7) because in the mapping from WS2 to WS3, the number of accessible terms increases by two.[7] Thus, SM/PM has an accessibility problem and this problem also holds for Late Merge, which requires a complex operation in addition to the accessibility problem. Therefore, SM, PM, and Late Merge are considered illegitimate.

Chomsky (2019a, b, 2020) and Chomsky, Gallego, and Ott (2019) assume that the structure-building system includes not only set-Merge but also pair-Merge. Pair-Merge is an asymmetric operation that yields adjunction structures.[8] Consider (12).

(12) [the $[_{NP}$ $[_{AP}$ young] $[_{NP}$ man]]]

The AP *young* adjoins to the NP *man* via pair-Merge, creating an adjunction

structure. Pair-Merge is asymmetric in the sense that the host NP *man* retains its property even after the adjunct AP *young* adjoins. That is, *man* remains an NP after *young* pair-merges with *man*. Pair-Merge forms the adjunction structure as in (12) and the syntactic object created by pair-Merge is also represented as the ordered pair <*young*, *man*>, which means that *young* adjoins to *man*.

We have reviewed the minimalist framework that we adopt. With this framework in mind, we discuss the NRCs in the following sections.

3. Previous Analyses of Non-Restrictive Relative Clauses and Their Problems

3.1. Subordinate Clause Analysis

Jackendoff (1977) is the first to argue that the NRC is a subordinate clause embedded in its antecedent. For example, the NRC in (13a) is the daughter of N‴ as shown in (13b).

(13) a. The flowers, which she planted in the front yard, are growing well.
b. [$_{N'''}$ the [$_{N''}$ [$_{N'}$ flowers]] [$_{S'}$ which she planted in the front yard]]

N‴ is a ternary branching structure consisting of *the*, N″, and S′ (the NRC). Because the NRC is contained in N‴, it is a subordinate clause. Let us call this analysis the subordinate-clause analysis (SCA). Later, the SCA is updated to a binary branching structure in which the NRC is right-adjoined to the antecedent DP, as indicated in (14) (e.g., Toribio (1992)).

(14) [$_{DP}$ [$_{DP}$ the [$_{NP}$ flowers]] [$_{CP}$ which she planted in the front yard]]

The noun phrase is DP, based on Abney's (1987) DP hypothesis, and the NRC right-adjoins to DP *the flowers*. One piece of empirical evidence for the SCA is the fact that the combination of the antecedent and NRC can occur in the focus position of a cleft and pseudo-cleft sentence. Consider (15).

(15) a. It was Fred, who you met at my party, that I was just talking to on the phone.
b. What Mary likes the most is dancing, which unfortunately doesn't appeal to Roger at all.

(McCawley (1998: 449))

(15a) shows that the antecedent *Fred* occupies the cleft-focus position along with the NRC. (15b) indicates that the antecedent *dancing* occurs with the NRC in the post-copular position, which is the focus position. Given that a syntactic constituent occupies the focus position, (15a, b) support the idea that the antecedent and the NRC form a constituent. However, there is an empirical problem with the SCA. The SCA predicts that a bound pronoun in the NRC can be bound by the matrix element; however, this prediction is incorrect.

(16) a. Everyone$_i$ spoke about the museum that he$_i$ had visited.
b. *Everyone$_i$ spoke about the Millennium Dome, which he$_i$ had visited.

(cf. De Vries (2006: 256))

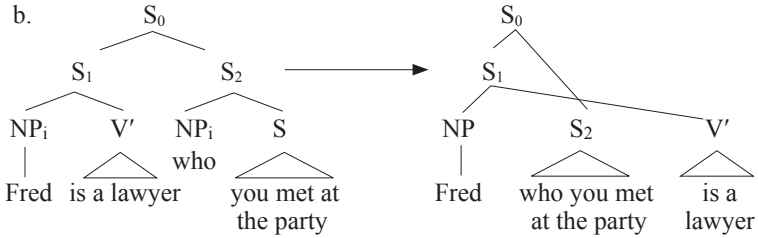
In (16a), the bound pronoun *he* in the RRC can refer to *everyone* in the matrix clause, whereas in (16b), the bound pronoun *he* in the NRC cannot. This contrast demonstrates that the bound pronoun in the NRC cannot be licensed by an antecedent in a matrix clause. If the NRC were a subordinate clause embedded in the antecedent DP, the bound pronoun would be bound by the antecedent; thus, (16b) should be grammatical, as in (16a). Therefore, the SCA cannot account for the ungrammaticality of (16b).

3.2. Main Clause Analysis

Ross (1986), Emonds (1979), McCawley (1998), among others, argue that NRCs are the main clauses. For example, McCawley (1998) analyzes NRCs, as in (17b).

(17) a. Fred, who you met at the party, is a lawyer.

b.



(McCawley (1998: 449))

The sentence in (17a) has the structure in (17b). The NRC (S_2) combines with S_1, forming S_0.[9] Note that the NRC does not form a constituent with the antecedent *Fred*. It moves to a position right-adjacent to *Fred* in the tree diagram. McCawley names this operation adposition, in contrast to adjunction by which an RRC forms a constituent with an antecedent. Adposition is a type of movement in which an element is moved to a position adjacent to its target but does not form a syntactic constituent with it. Thus, adposition enables the NRC to immediately follow its antecedent without forming a constituent, as shown in (17b). We call the analysis in (17b) the main-clause analysis (MCA). The MCA can solve the empirical problem with the SCA. Consider (16), which is repeated here as (18).

(18) a. Everyone$_i$ spoke about the museum that he$_i$ had visited.
b. *Everyone$_i$ spoke about the Millennium Dome, which he$_i$ had visited.
(cf. De Vries (2006: 256))

As we have seen, *he* in (18a) can take *everyone* as its antecedent, whereas *he* in (18b) cannot. Although the SCA cannot account for the ungrammaticality of (18b), the MCA can provide a straightforward account. Under the MCA, (18b) has the structure of (19) (irrelevant details are omitted).

(19) [$_{CP0}$ [$_{CP1}$ everyone$_i$ spoke about the Millennium Dome], [$_{CP2}$ which he$_i$ had visited]][10]

CP_1 (*everyone spoke about the Millennium Dome*) and CP_2 (*which he had visited*) are the main clauses and are combined into CP_0. In this structure, the bound pronoun *he* is not c-commanded by *everyone*. Given that bound pronouns must be c-commanded by their antecedents, *he* has no licensor. Therefore, (18b) is ungrammatical. The MCA can also account for the grammaticality of (15a, b) (which is supporting evidence for the SCA) with recourse to adposition. For instance, (15a) is analyzed, as shown in (20).

(20) a. [$_{CP0}$ [$_{CP1}$ it was Fred that I was just talking to on the phone] [$_{CP2}$ who you met at my party]]
b. [$_{CP0}$ [$_{CP1}$ it was Fred [$_{CP2}$ who you met at my party] that I was just talking to on the phone]]

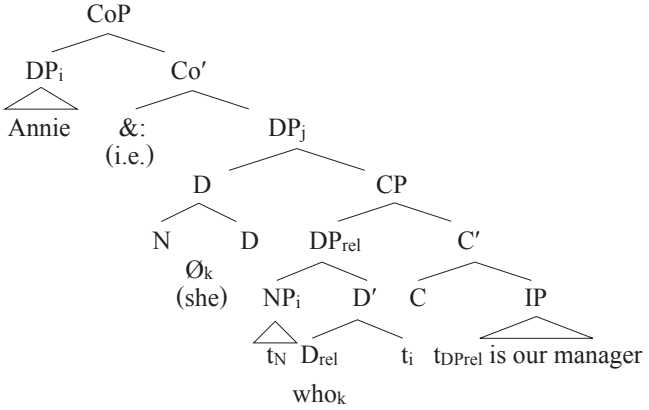
In (20a), the main clause CP_1 is combined with the NRC CP_2 into CP_0, and the NRC is in turn moved to a position adjacent to the antecedent *Fred* by adposition as in (20b), deriving the sentence in (15a). It is true that the MCA can account for the grammaticality of (15a, b) and solve the empirical problem with the SCA; however, it has a theoretical problem that the SCA does not face. The MCA invokes the construction-specific operation: adposition. Although it may be that if we posit adposition, we can account for the grammaticality of (15a, b), adposition is an operation specific to NRCs. Therefore, it is undesirable to stipulate this operation within the minimalist framework.

3.3. De Vries's (2006) Analysis

Let us introduce a Minimalist-based analysis of NRCs. De Vries (2006) argues that appositive relative clauses (NRCs) are semi-free relative clauses with a null pronominal head. For instance, the structure of (21a) is (21b).

(21) a. Annie, (i.e., she) who is our manager

b.



(De Vries (2006: 248))

The NRC antecedent *Annie* and the semi-free relative DP (the NRC) are coordinated by the coordinator *&:*, which means *namely*. This structure guarantees the appositive relation between the antecedent and the NRC. Let us examine the derivation of the second conjunct DP in detail. NP moves to Spec-DP_{rel} to establish an agreement relation with D_{rel}, which makes *who* co-indexed with *Ø*. DP_{rel} then undergoes movement to Spec-CP. The phonologically empty D head in turn takes CP as its complement. N in turn moves to the D head, deriving the second conjunct DP. Note that the complex head [N+D] is equivalent to a personal pronoun, which is represented as $Ø_k$. The analysis in (21b) is supported by the fact that the coordinator and the null D can overtly occur. Consider (22).

(22) a. Anne, or she who is our manager. (De Vries (2006: 244))

b. [[$_{DP}$ Annie] [[$_{\&:}$ or] [$_{DP}$ [$_{D}$ she] [$_{CP}$ who is our manager]]]]

We observe that *or* and *she* can overtly appear, as in (22a), showing that *&:* is realized as *or* and the D head is realized as *she* as in (22b). It is important to note that De Vries's analysis is a type of SCA because the NRC is embedded in DP. The analysis then has the problem that it cannot account for the ungrammaticality of (16b), repeated here as (23a).

(23) a. *Everyone$_i$ spoke about the Millennium Dome, which he$_i$ had visited.
(cf. De Vries (2006: 256))

b. [everyone$_i$ [spoke about [[$_{DP}$ the Millennium Dome] [&: [$_{DP}$ Ø [$_{CP}$ which he$_i$ had visited]]]]]]

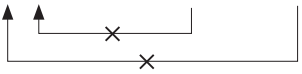
Under De Vries's analysis, *everyone* c-commands *he* because the NRC is embedded in DP, as shown in (23b). Therefore, the analysis incorrectly predicts that *everyone* can be the antecedent of *he*. De Vries (2006) circumvents this problem by suggesting that the second conjunct is invisible to c-command because the operation b-Merge is involved in the second conjunct. However, positing the new operation b-Merge is against the spirit of the Minimalist Program. Therefore, it is ideal that we account for the ungrammaticality of (23a) without recourse to b-Merge.

We have seen that there is a problem with Jackendoff's (1977) SCA, De Vries's (2006) SCA, and the MCA, which prompts us to explore another approach.

4. Proposal

To begin with, we again consider pair-Merge.

(24) [$_{DP}$ [$_{NP}$ [$_{AP}$ young] [$_{NP}$ men]]]



Pair-Merge is an operation yielding adjunction structures and pair-merged elements form an ordered pair <α, β>. In (24), *young* is pair-merged with *men*, creating the ordered pair <*young*, *men*>. One of the properties of pair-Merge is that the members of the ordered pair <α, β> formed by pair-Merge cannot undergo movement. In (24), neither *young* nor *men* can move. Thus, the pair-Merge property can be described by (25).

(25) The members of the ordered pair created by pair-Merge are inaccessible to Merge.[11]

Based on (25), we follow Kitahara (2019) and Tozawa (2020, 2022) and assume that the pair-merged elements are inaccessible to MERGE.[12] SM/PM does not then violate (7) as long as it is pair-Merge. Consider the derivation in (26), where the pair-MERGE by SM/PM occurs.

(26) a. WS1 = [a, b, c] (three accessible terms)
b. WS2 = [{a, b}, c] (four accessible terms)
c. WS3 = [{a, b}, <b, c>] (four accessible terms)

Suppose the derivation proceeds to WS2, where the number of accessible terms is four. Here, SM/PM is applied to WS2. That is, *b* in {*a*, *b*} undergoes SM with *c*, forming an ordered pair <*b*, *c*>. The number of accessible terms does not increase by more than one during the mapping from WS2 to WS3. The number of accessible terms in WS3 is four: *a*, *b*, {*a*, *b*}, and <*b*, *c*>. Note that *b* and *c* in <*b*, *c*> are inaccessible terms because they are members of an ordered pair. Then, a change from WS2 to WS3 observes (7). Therefore, we propose (27).

(27) SM/PM is permitted as long as it is pair-Merge.[13]

In the following sections, we analyze the NRCs based on this proposal and consider their properties.

5. Analysis

We argue that an NRC is coordinated with a main clause CP and the NRC antecedent undergoes SM/PM with the NRC. For example, the sentence in (28a) has the derivation in (28b-g).

(28) a. I telephoned Rod, who had called while I was out.

b. WS1 = [&, C, T, I, *v**, {telephoned, Rod}, {who had called while I was out}]

c. WS2 = [&, C, T, I, *v**, {telephoned, Rod}, <Rod, {who had called while I was out}>]

...

d. WS3 = [&, {C, {I, {T, {I, {*v**, {telephoned, Rod}}}}}}, <Rod, {who had called while I was out}>]

e. WS4 = [&, {{C, {I, {T, {I, {*v**, {telephoned, Rod}}}}}}, <Rod, {who had called while I was out}>}]

f. WS5 = [{&, {{C, {I, {T, {I, {*v**, {telephoned, Rod}}}}}}, <Rod, {who had called while I was out}>}}]

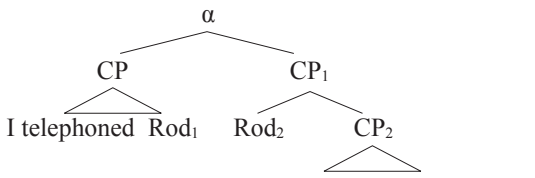
g. WS6 = [{{C, {I, {T, {I, {*v**, {telephoned, Rod}}}}}} {&, {{C, {I, {T, {I, {*v**, {telephoned, Rod}}}}}}, <Rod, {who had called while I was out}>}}}]

We assume that the derivation has reached WS1, where matrix VP and the NRC are formed. Pair-MERGE is applied to WS1, yielding WS2, where *Rod* adjoins to the NRC. Because the NRC is created independently from matrix VP, pair-Merge of *Rod* with the NRC is SM/PM. Given (27), SM/PM is possible.[14] Subsequently, MERGE is repeatedly applied so that the derivation reaches WS3, where the main clause CP is formed. MERGE changes WS3 into WS4, in which the main clause CP is merged with the NRC. Next, MERGE maps WS4 into WS5. In WS5, the phonologically null coordinator & is merged with the set consisting of two CPs (the main clause and the NRC). Finally, MERGE is applied to WS5, yielding WS6, where the main clause CP moves to the so-called specifier of &P.[15] In this way, the sentence including the NRC is derived.[16] We call this analysis the SM/PM analysis.

One might wonder how a copy relation is assigned between *Rod* in the main clause and *Rod* in the NRC. Following Chomsky (2023), we assume that the copy relation is established on the basis of a c-command relation. Note that in WS4, *Rod* in the NRC c-commands *Rod* in the main clause. For easier under-

standing, let us consider the tree diagram in (29).

(29)



Rod$_2$ adjoins to CP of the NRC, creating the adjunction structure. As a result, CP consists of two segments. If we assume the segment-category distinction (May (1985) and Chomsky (1986)), *Rod*$_2$ is not dominated by the CP of the NRC because it is not dominated by both of the two CP segments. Since the α that dominates *Rod*$_2$ dominates *Rod*$_1$, *Rod*$_2$ c-commands *Rod*$_1$. Based on the c-command relation, the copy relation between *Rod*$_1$ and *Rod*$_2$ is established. Here, the question arises of which copy is pronounced.[17] We argue that we can take *Rod*$_1$ to be the higher copy because the main clause including *Rod*$_1$ undergoes movement at the stage of WS6, occupying a higher position than the NRC. Given that the structurally higher copy is pronounced, it is *Rod*$_1$ that is pronounced. We also argue that the NRC immediately follows *Rod*$_1$ because the lower copy *Rod*$_2$ left-adjoins to the NRC.[18] The left-adjunction leads to the string *Rod*, *who had called while I was out*.[19] Here, we discuss whether our analysis can account for (15) and (18), which support the SCA and the MCA, respectively. First, consider (15a), repeated here as (30a).

(30) a. It was Fred, who you met at my party, that I was just talking to on the phone.
b. WS1 = [&, {it was Fred that I was just talking to on the phone}, {who you met at my party}]
c. WS2 = [&, {it was Fred that I was just talking to on the phone}, <Fred, {who you met at my party}>]
d. WS3 = [&, {{it was Fred that I was just talking to on the phone}, <Fred, {who you met at my party}>}]

e. WS4 = [{&, {{it was Fred that I was just talking to on the phone}, <Fred, {who you met at my party}>}}]

f. WS5 = [{{it was Fred that I was just talking to on the phone}, {&, {{it was Fred that I was just talking to on the phone}, <Fred, {who you met at my party}>}}}]

Suppose that the derivation proceeds to WS1, where the cleft sentence and the NRC are created. SM/PM applies to WS1, so that the antecedent *Fred* pair-merges with the NRC in WS2. MERGE is then repeatedly applied, yielding WS3, WS4, and WS5. In WS3, the main clause is combined with the NRC. In WS4, & is merged with the set of the main clause and the NRC. In WS5, the main clause undergoes movement. As *Fred* in the main clause occupies a higher position than *Fred* in the NRC, *Fred* in the main clause is pronounced.[20] Also note that this NRC immediately follows *Fred* because *Fred* left-adjoins to the NRC, as shown in WS5. That is, although the NRC is the second conjunct, it is pronounced not in the second conjunct position but in the position immediately after *Fred* in the main clause. Consequently, (30a) is derived. In this way the SM/PM analysis guarantees that NRCs immediately follow antecedents without recourse to adposition, which is an ad-hoc operation. Next, we consider (18), which is evidence for the MCA. (18b) is repeated here as (31a) and has the structure in (31b) under the SM/PM analysis.

(31) a. *Everyone$_i$ spoke about the Millennium Dome, which he$_i$ had visited.

b. WS = [{{everyone$_i$ spoke about the Millennium Dome}, {&, {{everyone$_i$ spoke about the Millennium Dome} <{the Millennium Dome}, {which he$_i$ had visited}>}}}]

The NRC antecedent *the Millennium Dome* in the main clause undergoes SM/PM with the NRC. The SM/PM is allowed because it is pair-Merge. After SM/PM of the antecedent with the NRC, the coordinate structure is formed. In this structure, the bound pronoun *he* is not c-commanded by the antecedent *everyone*. Thus, the bound pronoun is not licensed, resulting in the ungrammaticality

of (31a).[21, 22]

6. A Consequence: A Root Property of Non-Restrictive Relative Clauses

Let us briefly explore one consequence of our analysis. We argue that the main clause and NRC are coordinated. Therefore, our analysis is a kind of MCA. We then predict that NRCs behave like root sentences. This prediction is borne out. Consider (32).

(32) a. It may clear up, in which case would you mind hanging the washing out? (Huddleston and Pullum (2002: 1061))
b. *Mary knows the man by whom will we be forgotten for what we have done?

The contrast between (32a) and (32b) shows that NRCs allow Subject-Aux Inversion (SAI), whereas RRCs do not. Given that SAI occurs in root interrogative sentences, (32a) supports the claim that NRCs are main clauses.

7. Conclusion

In this study, we have argued that the NRC is coordinated with the main clause and the NRC antecedent is sideward-merged or parallel-merged with the NRC under the assumption that SM/PM is possible if it is pair-Merge.[23] We have shown that our analysis can solve the problems of the SCA and MCA. To the extent that the SM/PM analysis of NRCs is on the right track, SM/PM cannot be entirely eliminated from the theory of grammar.

Notes

1. Citko (2016) proposes that PM is involved in the derivation of NRCs. Our analysis differs from Citko's in some respects, as we will see in Section 5.
2. For convenience, we use the angled bracket notation to refer to the set of syntactic objects in WS.

3. For details on Remove, see Chomsky (2019a, 2021).
4. Although Nunes (2001) names a series of operations in (10) sideward movement, we call it SM because movement is unified into Merge in the recent minimalist framework.
5. There is no distinction between SM and PM under the MERGE-based derivation. Therefore, Merge of *b* with *c* can be regarded as SM or PM. Henceforth, we use the expression "SM/PM" because it does not make any difference whether Merge of *b* with *c* is SM or PM.
6. Neither *b* of {*a*, *b*} nor *b* of {*b*, *c*} c-commands the other. Therefore, MS cannot make one of the two copies inaccessible.
7. An anonymous reviewer notes that there is no violation of (7) if the derivation in (11) includes WS3′ between WS2 and WS3, as shown in (i).
   (i) a. WS1 = [a, b, c] (three accessible terms)
   b. WS2 = [{a, b}, c] (four accessible terms)
   c. WS3′ = [{a, b}, b, c] (five accessible terms)
   d. WS3 = [{a, b}, {b, c}] (six accessible terms)

   Suppose that the derivation has reached WS2. The copying operation changes WS2 into WS3′, where the copy of *b* is newly added. The number of accessible terms is five: *a*, *b*, *c*, *b* of {*a*, *b*}, and {*a*, *b*}. The mapping from WS2 to WS3′ observes (7) because the number of the accessible term increases by one. MERGE is then applied to WS3′, yielding MS3, where the set {*b*, *c*} is newly created. The number of accessible terms is six: *a*, *c*, *b* of {*a*, *b*}, *b* of {*b*, *c*}, {*a*, *b*}, and {*b*, *c*}. The mapping from WS3′ to WS3 satisfies (7) because the number of accessible terms increases by one. The derivation in (11) seems to be legitimate if WS3′ is added between WS2 and WS3. However, the copying operation cannot update WS. We have assumed in (4) that MERGE updates WS into WS′. It is then impossible that the copying operation updates WS. Therefore, WS3′ cannot be added between WS2 and WS3.
8. For details on pair-Merge, see Chomsky (2000, 2004).
9. NRC (S_2) is derived as follows: (i) the pronoun *he* co-referential with the NRC antecedent is replaced by the relative pronoun *who*, and (ii) *who* is moved to the NRC-initial position. Thus, the NRC in (17b) is derived.
10. We update the MCA by replacing S with CP in (19).
11. Given that Merge includes External Merge (EM) and Internal Merge (IM), the members of the ordered pair created by pair-Merge are inaccessible to both EM and IM.
12. Kitahara (2019) treats head-movement as a syntactic operation, arguing that it

does not violate Resource Restriction (RR), which is a predecessor of MY, because the elements of an ordered pair are inaccessible to MERGE. For example, the derivation of T-to-C movement is illustrated in (i), where EA represents an external argument.

(i) a. WS1 = [C, {EA, {T, *v*P}}] (six accessible terms)
b. WS2 = [<T, C> {EA, {T, *v*P}}] (six accessible terms)
c. WS3 = [{<T, C> {EA, {T, *v*P}}}] (seven accessible terms)

WS1 includes six accessible terms: C, EA, T, *v*P, {T, *v*P}, and {EA, {T,*v*P}}. MERGE applies to WS1, yielding WS2, in which the ordered pair <T, C> is newly formed. Importantly, neither T nor C of <T, C> is an accessible term. WS2 then contains six accessible terms: EA, T, *v*P, {T, *v*P}, {EA, {T, *v*P}}, and <T, C>. Therefore, the mapping from WS1 to WS2 observes RR. Next, MERGE applies to WS2, changing WS2 to WS3, where the ordered pair <T, C> is merged with {EA, {T, *v*P}}. WS3 includes seven accessible terms: EA, T, *v*P, {T, *v*P}, {EA, {T, *v*P}}, <T, C>, and {<T, C> {EA, {T, *v*P}}}. Therefore, the mapping from WS2 to WS3 satisfies RR. In this way, head-movement does not violate RR. Tozawa (2020) adopts Kitahara's idea that pair-merged elements are inaccessible to MERGE and argues that SM is permitted as long as it is pair-Merge. Based on this argument, Tozawa argues for Nunes's (2001) SM-based analysis of RRC adjunction, giving a third-factor-based account of the reconstruction asymmetry between RRCs and complement clauses. In this paper, we advance Kitahara's idea by arguing that SM/PM is possible as long as it is pair-Merge and that NRC antecedents undergo SM/PM with NRCs.

13. Note that the element that undergoes set-Merge by PM remains accessible to MERGE, unlike the one that undergoes pair-Merge by PM. Citko (2005) argues that the parallel-merged element undergoes movement, which makes linearization possible. However, set-Merge by PM is impossible, as we have seen in Section 2; only pair-Merge by PM is possible. The element that undergoes pair-Merge by PM is not available for movement because it is inaccessible to MERGE.
14. In WS2 in (28), *Rod* left-adjoins to the NRC, forming the ordered pair <Rod, NRC>. As we have seen, the pair-merged elements are inaccessible to MREGE. One might wonder whether the elements in the NRC are also inaccessible to MERGE. We suggest that, as extraction from relative clauses is prohibited, the elements in the NRC are inaccessible to MERGE. The accessible terms are then decreased in the mapping from WS1 into WS2. We argue that the decrease of accessible terms is allowed because it does not violate (7). I thank an anonymous

reviewer for raising this issue.

15. We follow Chomsky's (2013) coordination analysis, under which the first and second conjuncts are merged, the set of conjuncts is combined with the coordinator, and the first conjunct is moved.
16. Citko (2016) argues that NRCs exist in Polish, suggesting that NRC antecedents undergo PM with NRCs. Under Citko's analysis, the antecedent undergoes PM with the NRC using Hornstein's (2009) Concatenate. Because Concatenate yields a structure without a label, the constituent consisting of the antecedent and the NRC has no label. The NRC moves to the matrix CP for the constituent to be labeled (for labeling, see Chomsky (2013)). Our analysis differs from Citko's in that the antecedent is not combined with the NRC by Concatenate and the NRC does not move to the matrix CP.
17. I thank two anonymous reviewers for bringing this issue to my attention.
18. Note that an NRC cannot precede its antecedent. Consider (i).

    (i) a. It rained all day yesterday, which I expected.
        b. *Which I expected, it rained all day yesterday.

    In (ia), the antecedent of the NRC is the main clause CP (*It rained all day yesterday*). (ib) shows that the NRC cannot precede the antecedent. We argue that the NRC must follow the antecedent for two reasons. One is that, as we have seen, the antecedent left-adjoins to the NRC. The other reason is that the NRC is the second conjunct. (ia) has the structure in (ii) (irrelevant details are omitted).

    (ii) [[$_{CP}$ it rained all day yesterday] [& [$_{CP}$ which I expected]]]

    We suggest that the NRC has to be the second conjunct because it gives the supplementary explanation of the antecedent. Since the second conjunct always follows the first conjunct, the NRC in (ib) cannot precede the antecedent. I thank an anonymous reviewer for raising the issue of the relation between the main clause and the NRC.
19. An anonymous reviewer brings up the following examples and points out that NRCs do not have to be adjacent to their antecedents.

    (i) a. Only the flower is used, which is not poisonous and is attached to the plant with a very fine stem. (Huddleston and Pullum (2002: 1066))
        b. Kim likes muffins$_i$, but Sandy prefers scones$_j$, which$_{i+j}$ they eat with jam. (Arnold (2007: 274))

    (ia) shows that the NRC is extraposed to sentence-final position, so that the NRC and its antecedent are separated. (ib) is an example of the split-antecedent NRC. The NRC does not immediately follow one of the antecedents, that is, *muffins*. To account for (ia, b), we follow Culicover and Rochemont (1990) in assuming that

extraposed elements are base-generated in the surface position without forming a constituent with their antecedents. We then suggest that the sentence in (ia) has the following structure.

(ii) [$_{CP}$ [$_{CP}$ only the flower is used] [$_{CP}$ which is not poisonous and is attached to the plant with a very fine stem]]

The NRC does not undergo SM/PM with the antecedent *only the flower*. Rather, it is base-generated in the surface position, that is, the CP adjunct position. We suggest that the NRC is related to its antecedent in the semantic component, although the relation is not established in syntax. In (ii), the antecedent does not left-adjoin to the NRC. Therefore, the NRC does not immediately follow the antecedent. We now turn to (ib). Based on Culicover and Rochemont's suggestion that the split-antecedent RRC is a case of extraposition, we suggest that the same analysis as in (ia) then holds for the split-antecedent NRC in (ib). The structure is shown in (iii).

(iii) [$_{CP}$ [$_{CP}$ Kim likes muffins$_i$, but Sandy prefers scones$_j$] [$_{CP}$ which$_{i+j}$ they eat with jam]]

The NRC is base-generated in the CP adjunct position and does not form a constituent with the antecedents *muffins* and *scones*. Therefore, the NRC is not adjacent to its antecedents.

20. I thank an anonymous reviewer for pointing out the problem of which copy of *Fred* is pronounced.
21. Condition C effects are observed in NRCs, as indicated in (i).

(i) *He$_i$ owns a car, which John$_i$ drives every day. (cf. De Vries (2006: 257))

Under our analysis, *he* does not c-command *John*. Therefore, we predict that (i) should be grammatical, contrary to fact. Following De Vries (2006), we suggest that Condition C is a discourse condition (perhaps a syntactic condition as well). We can say that (i) is ungrammatical for a discourse reason just like (ii).

(ii) He$_i$ owns a car. John$^*_i$ drives it every day. (De Vries (2006: 257))

Although in (ii) *he* does not c-command *John*, Condition C is induced as a discourse condition. Therefore, *he* cannot refer to *John* across a sentence boundary. I thank an anonymous reviewer for pointing out this issue.

22. Dobashi (2020) provides a principled account of the c-command invisibility of the bound pronoun in the NRC based on the concept of the termination of a derivation. According to Dobashi, the NRC is a terminated derivation, which is not syntactically integrated into the associated structure. Therefore, it is not merged with its antecedent in WS. Consequently, the bound pronoun in the NRC is not c-commanded by any element in the matrix clause. See Dobashi (2020) for details.
23. An anonymous reviewer notes that our analysis predicts that NRCs are observed

in all languages because pair-Merge, which is a general syntactic operation, is involved in the derivation of NRCs. However, this prediction seems incorrect, because Chinese lacks English-type NRCs, as demonstrated by Del Gobbo (2010). Citko (2016) alludes to the possibility that Chinese lacks English-type NRCs because it lacks the DP projection to which NRCs attach (see Cheng and Sybesma (1999) for the argument that Chinese has no DP). Citko eventually dismisses this possibility, but we agree with it, suggesting that in Chinese there is no DP that is combined with NRCs through pair-Merge and therefore that Chinese lacks English-type NRCs.

References

Abney, Steven Paul (1987) *The English Noun Phrase in Its Sentential Aspect*. Doctoral dissertation. MIT.

Arnold, Doug (2007) "Non-Restrictive Relatives Are Not Orphans." *Journal of Linguistics* 43: 271-309.

Cheng, Lisa Lai-Shen and Rint Sybesma (1999) "Bare and Not-So-Bare Nouns and the Structure of NP." *Linguistic Inquiry* 30: 509-542.

Chomsky, Noam (1986) *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000) "Minimalist Inquiries: The Framework." In Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam (2004) "Beyond Explanatory Adequacy." In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, Volume 3, 104-131. Oxford: Oxford University Press.

Chomsky, Noam (2013) "Problems of Projection." *Lingua* 130: 33-49.

Chomsky, Noam (2019a) "Some Puzzling Foundational Issues: The Reading Program." *Catalan Journal of Linguistics. Special Issue, Generative Syntax: Questions, Crossroads, and Challenges*, 2019: 263-285.

Chomsky, Noam (2019b) MIT Lectures. April 10th and 12th.

Chomsky, Noam (2020) "The UCLA Lectures."
<https://ling.auf.net/lingbuzz/005485>

Chomsky, Noam (2021) "Minimalism: Where Are We Now, and Where Can We Hope to Go." *Gengo Kenkyu* 160: 1-41.

Chomsky, Noam (2023) "The Miracle Creed and SMT." Ms., University of Arizona and MIT.

Chomsky, Noam, Ángel J. Gallego, and Dennis Ott (2019) "Generative Grammar and

the Faculty of Language: Insights, Questions, and Challenges." *Catalan Journal of Linguistics. Special Issue, Generative Syntax: Questions, Crossroads, and Challenges*, 2019: 229-261.

Citko, Barbara (2005) "On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge." *Linguistic Inquiry* 36: 475-496.

Citko, Barbara (2016) "Types of Appositive Relative Clauses in Polish." *Studies in Polish Linguistics* 11: 85-110.

Culicover, Peter W. and Michael S. Rochemont (1990) "Extraposition and the Complement Principle." *Linguistic Inquiry* 21: 23-47.

Del Gobbo, Francesca (2010) "On Chinese Appositive Relative Clauses." *Journal of East Asian Linguistics* 19: 385-417.

Dobashi, Yoshihito (2020) *Externalization: Phonological Interpretations of Syntactic Objects*. London: Routledge.

Emonds, Joseph (1979) "Appositive Relatives Have No Properties." *Linguistic Inquiry* 10: 211-243.

Hornstein, Norbert (2009) *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.

Huddleston, Rodney and Geoffrey K. Pullum (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Jackendoff, Ray (1977) *X′ Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kitahara, Hisatsugu (2019) "(Head-)Adjunction by Pair-MERGE." paper presented at the workshop "Pair-Merge under MERGE" in the 12th International Spring Forum of the English Linguistic Society of Japan.

Lebeaux, David (1991) "Relative Clauses, Licensing, and the Nature of the Derivation." In Susan D. Rothstein (ed.), *Syntax and Semantics 25: Perspectives on Phrase Structure: Heads and Licensing*, 209-239. New York: Academic Press.

May, Robert (1985) *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

McCawley, James D. (1998) *The Syntactic Phenomena of English*, 2nd ed. Chicago: University of Chicago Press.

Nunes, Jairo (2001) "Sideward Movement." *Linguistic Inquiry* 32: 303-344.

Ross, John Robert (1986) *Infinite Syntax!*. Norwood, New Jersey: Ablex.

Toribio, Almeida J. (1992) "Proper Government in Spanish Subject Relativization." *Probus* 4: 291-304.

Tozawa, Takahiro (2020) "The Sideward Pair-MERGE of the Relative Clause and Its Antecedent." *Explorations in English Linguistics* 34: 71-97.

Tozawa, Takahiro (2022) “Hiseigen Kankeisetsu no Yokohoko Tsuiheigo Bunseki (A Sideward Pair-Merge Analysis of Non-restrictive Relative Clauses).” In Etsuro Shima, Naoto Tomizawa, Yoshiki Ogawa, Yoshihito Dobashi, Yosuke Sato, and Cornelia Daniela Lupşă (eds.), *Kotoba No Yoso* (Aspects of Language: An Interface between Present and Future), 294-305. Tokyo: Kaitakusha.
Vries, Mark de (2006) “The Syntax of Appositive Relativization: On Specifying Coordination, False Free Relatives, and Promotion.” *Linguistic Inquiry* 37: 229-270.
(Kitami Institute of Technology)
tozawata@mail.kitami-it.ac.jp