



that-Trace Effect in English: Box System and Labeling*

Ick-Hee Ihm (Hankuk University of Foreign Studies) · Haewon Jeon (Silla University) · Ji-Yeoun Lee (Gyeongsang National University)



This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: November 3, 2025

Revised: December 9, 2025

Accepted: December 20, 2025

Ick-Hee Ihm (First author)
Academic research professor,
Language Research Institute
Hankuk University of Foreign
Studies
Email: 2008764@hufs.ac.kr

Haewon Jeon (Co-author)
Visiting Associate Professor,
English Education Department
Silla University
Email: hwJeon@kmou.ac.kr

Ji-Yeoun Lee (Co-author)
Lecturer, Department of English
Language and Literature
Gyeongsang National University
Email: think2ce@naver.com

*This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5B5A16079970). We are grateful to two anonymous reviewers for their helpful comments and suggestions.

ABSTRACT

Ihm, Ick-Hee, Haewon Jeon and Ji-Yeoun Lee. 2025. *that*-Trace Effect in English: Box System and labeling. *Korean Journal of English Language and Linguistics* 25, 1704-1718.

The aim of this study is to account for the asymmetries between subject and object in English *wh*-movement, topicalization, and *wh*-in-situ within the framework of Chomsky's (2024) Box System (BS) and the Labeling Algorithm (LA) (Chomsky 2013, 2015). Building on recent developments in the Minimalist Program, the paper proposes that the (in)accessibility of a matrix phase head C to its goal determines the (un)grammaticality of a *wh*- or topic element extraction from an embedded clause. When the overt complementizer *that* inherits ϕ -features and a force feature (subsuming [Q] or [Topic]) to INFL, a <Q, Q> or <Topic, Topic> label is established within the embedded {XP, INFLP} configuration, rendering an embedded *wh*- or topic element inaccessible to the matrix C. Conversely, in the absence of the overt complementizer *that*, the embedded *wh*- and topic element is accessible to the matrix C.

KEYWORDS

Box System, *that*-trace effect, (in)accessibility, asymmetry, labeling

1. Introduction

that-trace effect refers to a phenomenon whereby a subject operator cannot be retrieved from a finite embedded clause generated by the overt complementizer *that*. Concrete examples are provided in (1a) and (1b) ((1a) is taken from Pesetsky (2017, p. 2), and (1b) taken from Lasnik and Saito (1993, p. 82)).

- (1) a. * Who do you think that *t* met Sue?
 b. * John, I think that *t* won the race.

We can see that, in (1a), extraction of the embedded subject operator (*who*), and in (1b), extraction of a subject (*John*) as a topicalization candidate are impossible in the presence of the overt complementizer *that*. By contrast, an object operator can be extracted from a finite embedded clause generated by the overt complementizer *that*. Concrete examples are given in (2a) and (2b) ((2a) is taken from Pesetsky (2017, p. 2), and (2b) taken from Lasnik and Saito (1993, p. 82)).

- (2) a. Who do you think that Sue met *t*?
 b. That race, I think that John won *t*.

We can see that, in (2a), a *wh*-object (*who*) is extracted, and in (2b), the object (*that race*) is topicalized despite the presence of the overt complementizer *that*.

However, this asymmetry between the extraction of a subject operator and that of an object operator is not seen when the overt complementizer *that* is absent, as shown in (3a), (3b), (4a), and (4b) ((3a) is taken from Pesetsky (2017, p. 2), and (3b) taken from Lasnik and Saito (1993, p. 82), and (4a) is taken from Pesetsky (2017, p. 2), and (4b) taken from Lasnik and Saito (1993, p. 82)).

- (3) a. Who do you think *t* saw Sue?
 b. John, I think *t* won the race.

- (4) a. Who do you think Sue met *t*?
 b. That race, I think John won *t*.

We can observe that there is no ban on extraction of the subject operators when the complementizer is phonologically null. On the other hand, in-situ *wh*-subject and *wh*-object show symmetry in acceptability regardless of the presence of the overt complementizer *that*, although the latter is somewhat marginal, as shown in (5) (taken from Lasnik and Saito (1993, p. 126)).

- (5) a. ? Who said (that) who bought the book?
 b. Who said (that) John bought what?

The aim of this article is to capture subject-object asymmetries that have been observed in English *wh*-movement, topicalization, *wh*-in-situ, based on a recent minimalist framework, Box Theory (Chomsky 2024), and Labeling Theory (Chomsky 2013, 2015). This article proposes that it is the (in)accessibility of a phase head *C* to a *wh*- and topic element that determines the (im)possibility of extracting the *wh*- or topic element from an embedded clause,

demonstrating that this proposal provides a simple and principled explanation to asymmetries between the subject operators and object operators in English. This article is structured as follows. Section 2 reviews a new version of Merge and recent theoretical frameworks in minimalist syntax: Box Theory. Section 3 proposes a solution to account for the asymmetries between the subject operators and object operators. Section 4 concludes.

2. Merge to MERGE and Chomsky’s (2024) Box System (BS)

This section reviews a new version of Merge and recent theoretical frameworks in minimalist syntax: Box Theory proposed by Chomsky (2024).

2.1 Merge to MERGE

Merge is the only structure building operation in human language assumed in minimalist syntax. Since its initial introduction in Chomsky (1995), Merge has long been assumed to perform two distinct tasks: (i) combining two syntactic objects (SOs) to create a new SO, and (ii) identifying and projecting a relevant label, as illustrated in (6). The label is notated by underline, and the symbol ‘ \rightarrow ’ reads as ‘returns.’

$$(6) \text{Merge}(x, y) \rightarrow \{\underline{x}, \{x, y\}\}$$

Chomsky (2013, 2015) redefines the role of Merge, separating the labeling task and proposing instead that labeling is performed by a Minimal Search (MS)-based independent Labeling Algorithm (LA) operation. Under this revision, Merge is responsible solely for combining two SOs, as shown in (7):

$$(7) \text{Merge}(x, y) \rightarrow \{x, y\}$$

In later work (Chomsky 2019a, 2019b, 2020), the concept of Merge is further revised to act not directly on individual SOs but on those within a workspace (WS).¹ This version of Merge is referred to as MERGE (in capital letters) and is defined as an operation that maps one WS to another, i.e., a new WS, as illustrated in (8).

$$(8) \text{WS}_1=[x, y, z] \rightarrow \text{MERGE}(x, y)^2 \\ \text{WS}_2=[\{x, y\}, z]$$

In (8), WS_1 comprises three SOs: x , y , and z . Any two of these SOs in WS_1 can be targeted by MERGE. Suppose MERGE targets x and y . This operation maps or updates WS_1 to WS_2 , where x and y form the set $\{x, y\}$. This type of set formation represents how conventional External Merge (EM) is implemented within the MERGE-based framework. Now, let us look at (9), where MERGE combines the set $\{x, y\}$ with one of its members, x term.³

$$(9) \text{WS}_1=[x, y, z] \rightarrow \text{MERGE}(x, y)$$

¹ Regarding Form Set, see Chomsky et al. (2023, p. 14) and Chomsky (2024b, p. 23).

² To distinguish WS from the set of SO, square brackets are used in the former, and curly braces in the latter.

³ In the $\text{WS}=[a, \{b, c\}]$, the constituents of the WS are a and $\{b, c\}$, while b and c are elements of the WS.

$$WS_2 = [\{x, y\}, z] \rightarrow \text{MERGE}(x, \{x, y\})$$

$$WS_3 = [\{x, \{x, y\}\}, z]$$

This operation maps or updates WS_2 to WS_3 , where x appears twice: once as a member of the set $\{x, \{x, y\}\}$, and again as a member of the subset $\{x, y\}$. This type of set formation represents how conventional Internal Merge (IM) is executed under the MERGE-based system. Chomsky (2024) notes that several constraints are imposed on this MERGE.⁴

2.2 Constraints on MERGE

Chomsky (2024) proposes that EM and IM are constrained by Principles [S] and [T] in (10).

(10) a. Principle [S]: The computational structure of language should adhere as closely as possible to SMT.
(Chomsky 2024, p. 19)

b. Principle [T]: All relations and structure-building operations are thought related, with semantic properties interpreted at CI.
(Chomsky 2024, p. 22)

Specifically, EM is constrained by dictation and requirement of Principles [S] and [T] in (11), and IM by those of Principles [S] and [T] in (12) ((11) and (12) are adapted from Nakashima (2025, p. 4)).

(11) a. Principle [S] dictates that input to EM must be picked out by Minimal Search (MS). For instance, P and Q are eligible to EM only if they are members of WS.⁵

b. Principle [T] requires that EM is associated with theta-related interpretation at CI, and X and Y are eligible to EM only if they are a pair of a theta-assigner and a theta assignee.

(12) a. Principle [S] dictates that MS limits the input of IM to a member of WS and its term.

b. Principle [T] dictates that only a member of a theta-structure qualifies for IM.

Principle [T] leads to the following entailments: (i) Once an element has undergone from a theta-position to a non-theta position through IM, it never moves out of there via IM (that is, there is no successive-cyclic movement in this framework), and (ii) EM creates an argument structure and IM creates force- and information-related structures with scopal properties. Chomsky (2024) names the former as a propositional domain and the latter as a clausal domain, and terms these two domains the Duality of Semantics principle (DS principle). We get the table 1 that shows how input and output of EM and IM are constrained by Principles [S] and [T].

Table 1. Input and Output Constraints of EM and IM due to Principles [S] and [T]

	EM	IM
Principle S: Minimal Search	Members of WS	A member of WS and its term
Principle T: Theta structure	Theta assigner and assignee	Theta assignee

⁴ Refer to Shim (2019, p. 107) regarding advantages of MERGE.

⁵ 1. MS is a third-factor principle that scans WS to find the closest, most relevant elements for a syntactic operation.

2. MS stops its search upon identifying the initial element within the c-command domain.

With these constraints on MERGE in place, Chomsky (2024) proposes Box Theory, which derives sentences with displacement in a markedly different manner from earlier minimalist frameworks such as Chomsky (2021) and Chomsky et al. (2023).

2.3 Chomsky’s (2024) Box System(BS)

In Chomsky (2024, p. 24), he notes that (i) it is unnecessary to differentiate between A- and A’-positions as well as A- and A’-movement. There’s just IM and EM, which are ineliminable, satisfying DS principle, and (ii) to adhere to DS principle, IM generates an element that does not engage with EM-generated structures within the propositional domain or with the procedures applicable therein. Then, in Chomsky (2024, p. 25), he notes that (i) the element E, which is IM-ed to the phase edge, exists in a box, isolated from the ongoing derivation D. Nonetheless, E must remain accessible to D at subsequent phase levels for interpretation at the interfaces. Additionally, there is no movement, and the boxed element is available for instructions at later phases. These operations are called Box System (BS). BS implies that (i) once a theta-marked *wh*-element undergoes IM to the lowest phase edge, its movement virtually terminates within narrow syntax, and (ii) successive-cyclic *wh*-movement is no more perceptible, so IM never fills [SPEC, CP] with the *wh*-syntactic object. Instead, C_[Q] accesses the *wh*-element to obtain feature instructions for scope interpretation at the Conceptual–Intentional (CI) interface and for *wh*-spell-out at the Sensory–Motor (SM) interface. To establish BS, Chomsky (2024) assumes the following points:

- (13) a. A verbal phase is v* and its internal arguments, and external argument enters in the next phase.
 b. A WH/Topic undergoes IM to SPEC-v*P.
 c. A higher phase-head locates an element at the phase-edge within a c-command configuration (cc-configuration) for interpretation at the interfaces.
 d. The SPEC of a phase will include boxed elements formed by IM but also an element merged by EM in a theta-position: predicate-internal subject for example, which can be raised to phase-internal SPEC-INFL (also its VO counterpart, object raising) (Chomsky 2024, p. 25). In other words, the phase edge (i.e., SPEC-v*) is accessible to C, regardless of whether it is the inner SPEC-v* created by IM (the ‘boxed’ position), or the outer SPEC-v* (i.e., predicate internal subject position associated with theta-marking).

As a reference, Chomsky (2023:b) classifies phases as in Table 2, to which relevant domains are added.

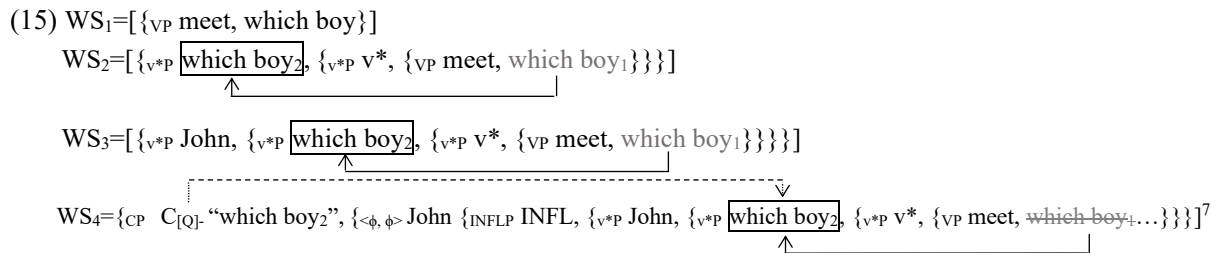
Table 2. Phases and Relevant Domains

1	[v* [2V NP ₂]]	v*-phase	NP ₂ is in a propositional domain
2	[_{v*P} NP ₁ [v* [2V NP ₂]] ↑	v*P-phase	NP ₁ is in a clausal domain
3	[_{v*P} subject [_{v*P} NP ₁ [v* [2V NP ₂]]]	Extended v*P phase	subject is in a propositional domain
4	[INFLP subject ₂ [_{v*P} subject ₁ [_{v*P} NP ₁ [v* [2V NP ₂]]]]]	non-phase	subject ₂ is in a propositional domain ⁶
5	[C [INFLP subject ₂ [_{v*P} subject ₁ [_{v*P} NP ₁ [v* [2V NP ₂]]]]]]]	C-phase	
6	[CP [C [INFLP subject ₂ [_{v*P} subject ₁ [_{v*P} NP ₁ [v* [2V NP ₂ ...]]]]]]]	CP-phase	

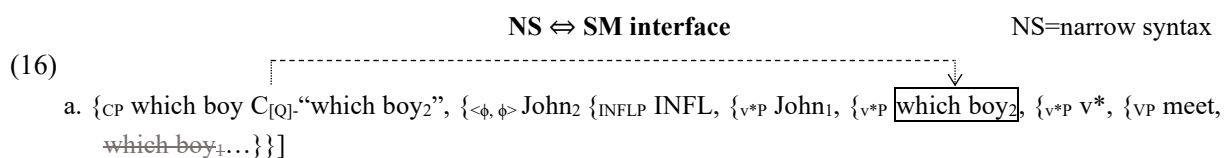
⁶ See Chomsky (2024, p. 32).

Within the BS framework and under the assumptions in (13), let us first consider how a simple question with a *wh*-object argument is derived, taking an example in (14). The sentence in (14) is assumed to undergo the derivation processes shown in (15).

(14) Which boy did John meet?



To describe (15) in detail, (i) WS₁: The members, a transitive verb *meet* and a *wh*-object *which boy*, undergo an external merge with each other, which generates a VP structure, where theta marking takes place (*which boy* is in a propositional domain) (WS₁).⁸ (ii) WS₂: A phase head v* externally merges with the VP, building a v*P structure, i.e., a verbal phase (v* and its internal arguments constitute a verbal phase), and the *wh*-object is obligatorily put in a box by the Duality of Semantics principle upon merging internally into the specifier of the phase v*P (henceforth, the element that is IMed to the phase edge is indicated by a square box. *which boy* is in a clausal domain). Form Copy (FC) then applies to the local pairs of identical inscriptions and the term *which boy₁* lacks phonological features due to FC (henceforth, we indicate FC in gray. See Kwon (2025, p. 181) for conditions for the operation of FC). (iii) WS₃: The subject *John* externally merge with the phase v*P, where theta marking takes place (the subject is in a propositional domain).⁹ (iv) WS₄: INFL externally merges with the v*P, followed by IM of the subject *John* to [SPEC, INFLP], and the SO is labeled <φ, φ> (Chomsky 2013, 2015). In the end, EM of C with <φ, φ>P closes the CP phase. At this phase level, C_[Q] accesses the boxed *which boy₂* for relevant feature instructions, obtaining the instructions (a. At this stage, for instance, an interpretable Q-feature of C (probe) and an uninterpretable Q-feature of *which boy* (goal) are valued. b. Henceforth, access is indicated by a dotted line). The C with relevant feature instructions, i.e., feature bundle, of *which boy* is indicated by C_[Q] “which boy₂”, where the subscript numeral is assigned just for expository purposes. It is crucial to recognize that, due to the boxed element’s immunity to Merge, only the valued properties of the *wh*-object should be present on C; the *wh*-object itself does not need to manifest in the [SPEC, CP] in the form of SO generated by IM. This syntactic representation brings about SM configuration as in (16).



⁷ In accordance with Chomsky’s (2021) claim that Tense is a feature of small v rather than INFL, we choose the term INFL instead of T.

⁸ Chomsky (2021) dictates that language must provide argument structure at the Conceptual-Intentional (CI) interface and Labeling Theory, and (Chomsky 2013, 2015) dictates that language must provide labeled SOs at the CI interface.

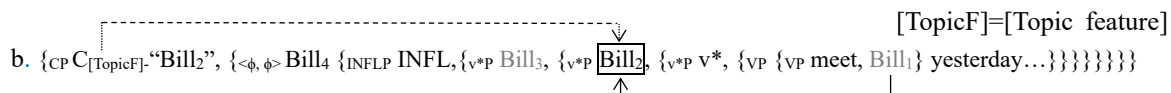
⁹ Chomsky (2024, p. 34: fn. 31) stipulates that the inner SPEC-v* created by IM belongs to the v*P phase, whereas the outer SPEC-v* created by EM of EA enters into the next phase.

- b. $\{_{CP} C_{[Q]} \text{“which boy}_2\text{”}, \{<\phi, \phi> John_2 \{_{INFLP} INFL, \{_{v^*P} John_1, \{_{v^*P} \text{which boy}_2, \{_{v^*P} v^*, \{_{VP} meet, \text{which boy}_{1...}\}}\}}\}}\}$ (throughout the paper, deletion is indicated by strikethrough)

Externalization is activated, after which the following processes are executed: (i) (16a): Given that $C_{[Q]}$ consists of feature bundle obtained through access to the boxed $which\ boy_2$, the phonological features of “which boy₂” must be spelled out at [SPEC, CP]. (ii) (16b): $C_{[Q]}$ “which boy₂” and $which\ boy_{2,1}$ are structurally identical and in a cc-configuration. Accordingly, the lower copies, $which\ boy_{2,1}$, are deleted. Likewise, $John_2$ and $John_1$ are structurally identical and in a cc-configuration, and the lower copy, $John_1$, is deleted by a universal economy principle.¹⁰ This series of narrow syntax and externalization operations yields a simple interrogative phrase with a *wh*-object argument without presupposing successive cyclic *wh*-movement.

Secondly, let us take on a case (17a) in which one same noun phrase form functions as a topicalized object on the one hand and as a subject on the other, simply put, a topicalized object and a subject are repetitions at CI, taking an example in (17). Let us suppose that derivation has reached the final WS stage (17b).

- (17) a. Bill, Bill met yesterday.



(17b) undergoes the same derivation processes as (WS₄ of (15)), except C’s topic feature. This syntactic representation yields SM configuration as in (18).

NS ⇔ SM interface

- (18)
- a. $\{_{CP} Bill C_{[TopicF]} \text{“Bill}_2\text{”}, \{<\phi, \phi> Bill_4 \{_{INFLP} INFL, \{_{v^*P} Bill_3, \{_{v^*P} \text{Bill}_2, \{_{v^*P} v^*, \{_{VP} \{_{VP} meet, \text{Bill}_1\} yesterday...}\}}\}}\}}\}$
- b. $\{_{CP} Bill C_{[TopicF]} \text{“Bill}_2\text{”}, \{<\phi, \phi> Bill_4 \{_{INFLP} INFL, \{_{v^*P} \text{Bill}_3, \{_{v^*P} \text{Bill}_2, \{_{v^*P} v^*, \{_{VP} \{_{VP} meet, \text{Bill}_1\} yesterday...}\}}\}}\}}\}$

Externalization is activated, after which the following processes are executed: (i) (18a): Given that $C_{[Q]}$ consists of feature bundle obtained through access to the boxed $Bill_2$, the phonological features of “Bill₂” must be spelled out at [SPEC, CP]. (ii) (18b): “Bill₂” and $Bill_{4,3,2,1}$ are in a cc-configuration. However, while “Bill₂” and $Bill_{2,1}$ are structurally identical copies, “Bill₂” and $Bill_4$ are not structurally identical copies because the former bears a Topic feature, whereas the latter does not, and hence “Bill₂” and $Bill_4$ are not semantically identical. $Bill_4$ is structurally identical to $Bill_3$. To put it differently, “Bill₂” and $Bill_4$ are not copy relations; they are repetition relations, which means that the two occurrences of *Bill* in (18a) are repetitions, not copies. Consequently, the lower copies— $Bill_{3,2,1}$ —undergo deletion by a universal economy principle. This series of narrow syntax and externalization operations generates the sentence (17) without assuming successive cyclic topic element-movement.

¹⁰ 1. Chomsky (2024, p. 35) defines a universal economy principle as deleting features that yield Externalization.
 2. Chomsky (2021, 2024) argues that *structurally identical* elements that are in a cc-configuration are interpreted as identical copies at the CI interface, while the lower copies of identical elements are deleted at the SM interface by a universal economy based rule under Externalization.
 3. Chomsky (2023b) states that X and Y are structurally identical if they have identical formal and semantic features while being non-distinct in phonological and morphological features. Structurally identical elements could indicate either repetition or copy relations. Copy relations are formed between two identical elements when they exist in the same phase, and all other such inscription pairs are repetitions (see Chomsky 2021, 2024).

Thirdly, let us consider how a simple question sentence with a *wh*-subject argument is derived, taking an example in (19). The sentence in (19) is assumed to undergo the derivational processes shown in (20).

(19) Who read the book?

(20) $WS_1 = [\{v^*P \text{ who } v^*, \{VP \text{ read, the book}\}\}]$
 $WS_2 = [\{<\phi, \phi> \text{ who}_2 \{INFLP \text{ INFL, } \{v^*P \text{ who}_1 v^*, \{VP \text{ read, the book}\}\}\}\}]$
 $WS_3 =$
 $\{CP \text{ } C_{[Q]} \text{ "who}_2", \{<\phi, \phi> \text{ who}_2 \{INFLP \text{ INFL, } \{v^*P \text{ who}_1 v^*, \{VP \text{ read, the book}\}\}\}\}\}$
 $WS_4 =$
 $\{CP \text{ } C_{[Q]} \text{ "who}_2", \{<\phi, \phi> \text{ who}_2 \{INFLP \text{ INFL, } \{v^*P \text{ who}_1 v^*, \{VP \text{ read, the book}\}\}\}\}\}$

To describe (20) in detail, (i) WS_1 : The subject *who* externally merge with the phase v^*P , where theta marking takes place. (ii) WS_2 : INFL externally merges with the v^*P , followed by IM of the subject *who* to SPEC-INFLP. There is no further *wh*-movement due to the Duality of Semantics principle. ‘No further movement’ brings about the effect of the vacuous movement hypothesis (VMH). Form Copy (FC) then applies to the local pairs of identical inscriptions. (iii) WS_3 : In the end, EM of C with $<\phi, \phi>P$ closes the CP phase. At this phase level, in accordance with (13d) assumption, $C_{[Q]}$ first finds the SPEC- v^*P for relevant feature instructions. But this position lacks phonological features due to FC. $C_{[Q]}$ then looks for the copy of the SPEC- v^*P that has the full instructions. That’s SPEC-INFLP. $C_{[Q]}$ accesses SPEC-INFLP for relevant feature instructions, obtaining the instructions.¹¹ This syntactic representation yields SM configuration as in (21).

NS \Leftrightarrow SM interface

(21) $\{CP \text{ who } C_{[Q]} \text{ "who}_2", \{<\phi, \phi> \text{ who}_2 \{INFLP \text{ INFL, } \{v^*P \text{ who}_1 v^*, \{VP \text{ read, the book}\}\}\}\}\}$

Externalization is activated, after which the following processes are executed: (i) Given that $C_{[Q]}$ consists of feature bundle obtained through access to the who_2 , the phonological features of “ who_2 ” must be spelled out at [SPEC, CP]. (ii) “ who_2 ” and $who_{2,1}$ are in a cc-configuration and structurally identical. The lower copy, $who_{2,1}$ are deleted by a universal economy principle, and (ii) This series of narrow syntax and externalization operations produces the sentence (19) without assuming successive cyclic *wh*-movement. Now, within this BS framework and under the assumptions in (13), let us delve into what derives the asymmetry between a subject operator and an object operator.

3. Proposal: (In)accessibility of C to Its Goal

We propose that accessibility of C to its goal yields a well-formed sentence, while inaccessibility of C to its goal

¹¹ Chomsky (2023b) suggests that at each phase, the phase head consults whatever is in the lowest phase edge, whether it is a boxed element in the clausal domain boxed or an externally merged predicate internal subject in the propositional domain. For further details, see *Theoretical Linguistics at Keio-EMU: Linguistics as Scientific Inquiry, Lecture 4* (Chomsky, 2023b), available at <https://www.youtube.com/watch?v=6yfNTQbpt8o&t=1383s>.

results in an ill-formed one. More specifically, If INFL inherited from $C(\textit{that})$ agrees with its goal (XP), and $\langle Q, Q \rangle / \langle \textit{Topic}, \textit{Topic} \rangle$ becomes a label of a $SO\{XP, INFLP\}$, then C at a matrix CP need not access XP to obtain its relevant features because a $\langle Q, Q \rangle / \langle \textit{Topic}, \textit{Topic} \rangle$ label is already determined within the $SO\{XP, INFLP\}$. This inaccessibility condition prevents a subject operator from being extracted out of the embedded clause. In contrast, if $\langle Q, Q \rangle / \langle \textit{Topic}, \textit{Topic} \rangle$ (depending on a sentence type) is not the label of the $SO\{XP, INFLP\}$, C at a matrix CP accesses its goal to obtain its relevant features. Based on this proposal, we will account for the fundamental cause of the asymmetry between a subject operator and an object operator, beginning with (1a).

3.1 (1a) and (1b)

Let us revisit (1a), repeated in (22). The derivation processes of (1a) are outlined in (23), and let us suppose that the derivation has reached the stage WS_1 .

(22) * Who do you think that t met Sue?

(23) $WS_1 = [\{\gamma(=?)\ \textit{who}_1, \{\beta(=v^*P)\ v^*, \{\alpha(=VP)\ \textit{meet}, \textit{Sue}\}\}\}]$
 $WS_2 = [\{C\textit{-that}, \{\delta(=INFLP)\ INFL, \{\gamma(=?)\ \textit{who}_1, \{\beta(=v^*P)\ v^*, \{\alpha(=VP)\ \textit{meet}, \textit{Sue}\}\}\}\}\}]$
 $WS_3 = [\{C\textit{-that}, \{\epsilon(=\langle Q, Q \rangle)\ \textit{who}_2\ \{\delta(=INFLP)\ INFL, \{\gamma(=v^*P)\ \textit{who}_1, \{\beta(=v^*P)\ v^*, \{\alpha(=VP)\ \textit{meet}, \textit{Sue}\}\}\}\}\}\}]$
 $WS_4 = [\{\zeta(=CP)\ C\textit{-that}, \{\epsilon(=\langle Q, Q \rangle)\ \textit{who}_2\ \{\delta(=INFLP)\ INFL, \{\gamma(=v^*P)\ \textit{who}_1, \{\beta(=v^*P)\ v^*, \{\alpha(=VP)\ \textit{meet}, \textit{Sue}\}\}\}\}\}\}\}]$
 $WS_5 = \{_{CP}\ C_{[Q]}\textit{-"who}_2", \dots \{\zeta(=CP)\ C\textit{-that}, \{\epsilon(=\langle Q, Q \rangle)\ \textit{who}_2\ \{\delta(=INFLP)\ INFL, \{\gamma(=v^*P)\ \textit{who}_1, \{\beta(=v^*P)\ v^*, \{\alpha(=VP)\ \textit{meet}, \textit{Sue}\}\}\}\}\}\}\}$
 $WS_6 = \{_{CP}\ C_{[Q]}\textit{-"who}_2", \dots \{\zeta(=CP)\ C\textit{-that}, \{\epsilon(=\langle Q, Q \rangle)\ \textit{who}_2\ \{\delta(=INFLP)\ INFL, \{\gamma(=v^*P)\ \textit{who}_1, \{\beta(=v^*P)\ v^*, \{\alpha(=VP)\ \textit{meet}, \textit{Sue}\}\}\}\}\}\}\}$

To describe (23) in detail, (i) WS_1 : The members, a transitive verb *meet* and an object *Sue*, undergo an external merge with each other, resulting in a VP structure where theta marking occurs. Then, a phase head v^* externally merges with the VP, yielding a v^*P structure. Subsequently, a *wh*-subject *who* bearing [uninterpretable Q] and [ϕ] features externally merge with the phase v^*P , where theta marking takes place. Incidentally, in this configuration, i.e., DP- v^*P configuration, if the *wh*-subject does not undergo raising, the labeling of γ becomes impossible because v^* has neither a Q feature nor an interrogative feature. So, labeling for the SO γ will be delayed until the *wh*-subject undergoes raising. (ii) WS_2 : INFL externally merges with the v^*P , followed by C (here, the overt complementizer *that*) externally merging with the SO δ . (iii) WS_3 : Prior to discussing this stage, let us look at the (24) (taken from Chomsky 2013: (14')).

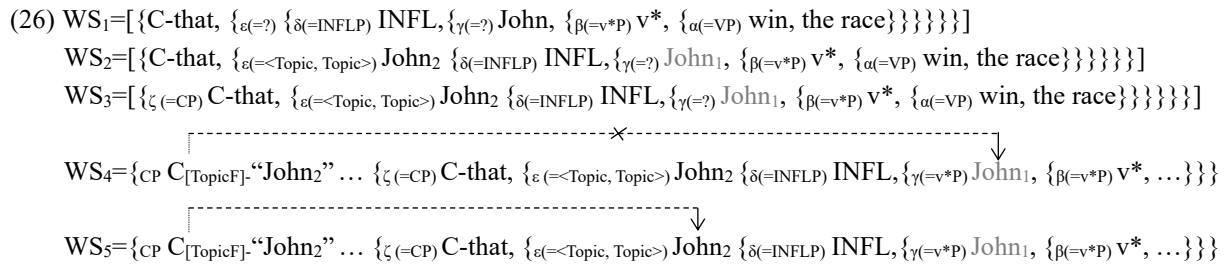
(24) * How many mechanics did they say that t fixed the cars?

With regard to (24), Chomsky makes the following analysis: INFL inherits ϕ -features and force feature (subsuming Q as a special case) from the $C(\textit{that})$, and force feature (subsuming Q as a special case) on INFL agrees with an uninterpretable Q-feature of a *wh*-subject (see Romaine 2015, p. 141-142) for evidence that *that* bears ϕ -

features).¹² Keeping these points in mind, let us revisit WS₃. The derivation contains both an overt complementizer *that* and a *wh*-subject. The C inherits its ϕ -features and force feature (subsuming Q as a special case) to INFL. The *wh*-subject raises to [SPEC, INFL], strengthening a weak T(ϕ -agreement checking occurs), and an uninterpretable Q-feature of *who* agrees with a force feature (subsuming an interpretable Q-feature) of INFL. Accordingly, the SO $\varepsilon, \{who, INFLP\}$, is labeled $\langle Q, Q \rangle$ (Chomsky 2013). FC then applies to the local pairs of identical inscriptions. At last, since the lower copy *who*₁ is invisible in MS, the merged SO γ is labeled v*P. (iv) WS₄: Since *that* and $\langle Q, Q \rangle$ are a $\{H, XP\}$ configuration, the SO ζ is labeled CP.¹³ (v) WS₅: Subsequently in the derivation, the phase head C_[Q] is introduced and the CP phase completes. At this phase level, in accordance with (13d) assumption, C_[Q] first finds the lowest [SPEC, v*P] for relevant feature instructions. But this position lacks phonological features due to FC. (vi) WS₆: C_[Q] then searches for the copy of *who*₁ that bears the full instructions. That turns out to be [SPEC, INFLP]. Accordingly, C_[Q] is expected to attempt to access [SPEC, INFLP] for relevant feature instructions. However, C_[Q] does not have motivation to access *who*₂ to obtain its relevant features because a $\langle Q, Q \rangle$ label—i.e., a QP label—is already determined in the ε . This condition of inaccessibility of C to *who*₂ means that *who* cannot be extracted from an embedded clause. Upon closer examination, it is ultimately the overt complementizer *that* that prevents C_[Q] from accessing *who*₂. Succinctly, the so-called blocking effect of the overt complementizer *that* on the access of C_[Q] to *who*₂ causes the ungrammaticality of (22).

Let us revisit (1b), repeated in (25). The derivation of (1b) proceeds in the same manner as that of (1a). The derivation processes of (1b) are outlined in (26), and let us suppose that the derivation has reached the stage WS₁.

(25) * John, I think that *t* won the race.



(i) WS₁: INFL externally merges with the v*P, followed by C (here, the overt complementizer *that*) externally merging with the SO ε . (ii) WS₂: Albuhayri (2019) argues that, in (27) below, [an uninterpretable Topic feature] originates on C and is inherited by INFL, and to check off [an uninterpretable Topic feature] of INFL, *flights to Chicago* moves to [SPEC, INFLP] under the complementizer *that* (taken from Albuhayri 2019: (66)).

(27) Bill warned us that [flights to Chicago we should try to avoid].

Inspired by the Albuhayri's (2019) argument, we propose that C (*that*) not only inherits ϕ -features to INFL but also inherits a force feature (subsuming a topic feature) to INFL. Keeping this in mind, let us return to WS₂. The subject *John* raises to [SPEC, INFLP], strengthening a weak T(ϕ -agreement checking occurs), and an interpretable topic-feature of *John* agrees with a force feature (subsuming an uninterpretable topic-feature) of INFL. The

¹² Refer to Shim (2019, p. 107) regarding features of *wh*-phrase and C.

¹³ Chomsky (2013, p. 47: fn. 47) claims that ϕ -feature inheritance from C to T has to leave [ϕ] on C intact.

embedded clause contains a topic-subject. Accordingly, the SO ε , {John, INFLP}, is labeled <Topic, Topic>. FC then applies to the local pairs of identical inscriptions. At last, since the lower copy *John*₁ is invisible in MS, the merged SO γ is labeled v*P. (iii) WS₃: Since *that* and <Topic, Topic> are a {H, XP} configuration, the SO ζ is labeled CP. (iv) WS₄: Later in the derivation, the phase head C_[TopicF] is introduced and the CP phase completes. At this phase level, in accordance with (13d) assumption, C_[TopicF] first finds the lowest [SPEC, v*P] for relevant feature instructions. But this position lacks phonological features due to FC. (vi) WS₅: C_[TopicF] then searches for the copy of *John*₁ that bears the full instructions. That turns out to be [SPEC, INFLP]. Accordingly, C_[TopicF] is expected to attempt to access [SPEC, INFLP] for relevant feature instructions. However, C_[TopicF] need not access *John*₂ to obtain its relevant features because a <Topic, Topic> label–i.e., a TopicP label–is determined in the SO ε . This unnecessary condition on the access of C_[TopicF] to *John*₂ prevents *John*₂ being extracted from the embedded clause. Upon closer examination, it is ultimately the overt complementizer *that* that prevents C_[TopicF] from accessing *John*₂. Succinctly, the so-called blocking effect of the overt complementizer *that* on the access of C_[TopicF] to *John*₂ results in the ungrammaticality of (25).

3.2 (2a) and (2b)

Let us revisit (2a), repeated in (28). The derivation processes of (2a) are outlined in (29), and let us suppose that the derivation has reached the stage WS₁.

(28) Who do you think that Sue met *t*?

(29) WS₁=[_{CP} C-that, {< ϕ , ϕ > Sue₂ {INFLP INFL, {v*P Sue₁, {v*P who₂ v*, {VP meet, who₁}}}}}}]]

-----*

↓

WS₂=[_{CP} C-that, {< ϕ , ϕ > Sue₂ {INFLP INFL, {v*P Sue₁, {v*P who₂ v*, {VP meet, who₁}}}}}}]]

-----*

↓

WS₃={_{CP} C_[Q]-“who₂”, ... {_{CP} C-that, {< ϕ , ϕ > Sue₂ {INFLP INFL, {v*P Sue₁, {v*P who₂ v*, {VP meet, who₁ ...}}}}}}]

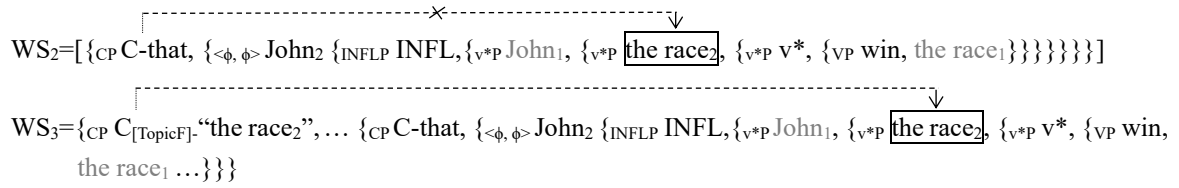
WS₂: Since the overt complementizer *that* does not bear Q-feature, it cannot enter into an Agree relation with the boxed *who*₂. WS₃: However, the C of the matrix CP bears a Q-feature, and thus it can enter into an Agree relation with the boxed *who*₂, rendering the boxed *who*₂ accessible to C_[Q]. The features of C and those of the boxed *who*₂ are then valued, with C assigning matrix scope to the boxed *who*₂. In contrast to *who* in (22=1a), the boxed *who*₂ does not enter into any agreement relation with the overt complementizer *that*, thereby permitting Agree between C_[Q] and the boxed *who*₂: the overt complementizer *that* has no impact on Agree between C_[Q] and the boxed *who*₂. When it comes to ϕ -agreement of an object, Shim’s (2020) analysis is considered the most valid to date.¹⁴

Let us next revisit (2b), repeated in (30). The derivation of (2b) proceeds in the same manner as that of (2a). The derivation processes of (2b) are outlined in (31), and let us assume that the derivation has reached the stage WS₁.

(30) That race, I think that John won *t*.

(31) WS₁=[_{CP} C-that, {< ϕ , ϕ > John₂ {INFLP INFL, {v*P John₁, {v*P the race₂, {v*P v*, {VP win, the race₁}}}}}}]]

¹⁴ We will explore this part more deeply in the future.

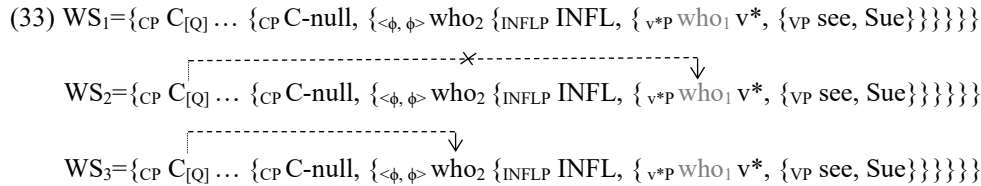


WS₂: Since the overt complementizer *that* does not bear a topic-feature, it cannot enter into an agreement relation with the boxed *the race*₂. WS₃: However, the C of the matrix CP bears a topic-feature and can thus enter into an Agree relation with the boxed *the race*₂, rendering *the race*₂ accessible to C_[TopicF]. The features of C and those of *the race*₂ are then valued, with C assigning matrix scope to *the race*₂. In contrast to the subject *John* in (1b), the boxed *the race*₂ does not enter into any agreement relations with the overt complementizer *that*, thereby permitting Agree between C_[TopicF] and the boxed *the race*₂: the overt complementizer *that* has no effect on Agree between C_[TopicF] and the boxed *the race*₂.

3.3 (3a) and (3b)

Let us revisit (3a), repeated in (32). The derivation processes of (3a) are outlined in (33), and let us suppose that the derivation has reached the stage WS₁.

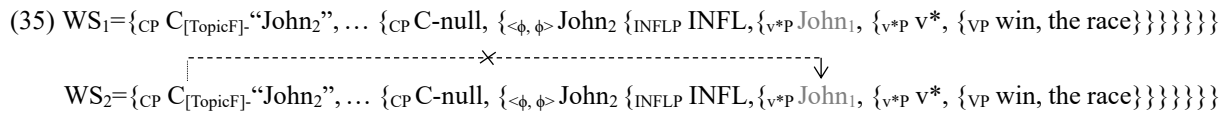
(32) Who do you think *t* saw Sue?



To elaborate on (33), (i) WS₁: The embedded C, lacking an overt complementizer, is de-phased, i.e., loses phasal status (see Chomsky 2015). Chomsky (2013) analyzes that when the complementizer is null, i.e., when ForceP is absent, INFL cannot inherit the feature F from C. This absence of feature inheritance opens a way to enable the matrix phase head C_[Q] to obtain relevant feature instructions from within the embedded clause. (ii) WS₂: In accordance with (13d) assumption, C_[Q] first finds the lowest [SPEC, v*P]. But this position lacks phonological features due to FC. C_[Q] then searches for the copy *who*₂ that bears the full instructions, which turns out to be [SPEC, INFLP]. C_[Q] accesses *who*₂ and terminates the search for its Agree target.

Let us revisit (3b), repeated in (34). The derivation of (3b) proceeds in the same manner as that of (3a). The derivation processes of (3b) are outlined in (35), and let us assume that the derivation has reached the stage WS₁.

(34) John, I think *t* won the race.



$$WS_3 = \{_{CP} C_{[TopicF]} \text{“John}_2\text{”, ... } \{_{CP} C\text{-null, } \{_{\langle \phi, \phi \rangle} John_2 \{_{INFLP} INFL, \{_{v^*P} John_1, \{_{v^*P} v^*, \{_{VP} win, the race\}}\}}\}}\}$$

As shown in WS_{1-3} , the absence of force-feature inheritance to INFL enables the matrix phase head $C_{[TopicF]}$ to access $John_2$ and obtain relevant feature instructions from within the embedded clause.

3.4 (5)

It was mentioned in the introduction section that in contrast to overt *wh*-movement, *wh*-in-situ in a subject and object is symmetric regardless of the presence of a complementizer, i.e., it does not exhibit a sharp contrast in acceptability between subject and object cases. Specifically, *wh*-in-situ may appear in the object position, as in (36a=5a), and in the subject position, as in (36b=5b), although the latter is somewhat marginal.

(36) a. Who said (that) John bought what?

b. ? Who said (that) who bought the book?

(Lasnik and Saito 1993, p. 126)

Let us examine how these sentences are derived. Let us suppose that the derivation of (36a) has reached (37), and that of (36b) has reached (38). At this stage, $C_{[Q]}$ will attempt to access the embedded clause *wh*-phrase: the boxed $what_2$ in (37) and the who_2 in (38).

$$(37) \quad WS = \{_{CP} C_{[Q]} \{_{\langle Q, Q \rangle} who_2 \{_{INFLP} INFL, \{_{v^*P} who_1, ... \{_{CP} C\text{-(that), } \{_{\langle \phi, \phi \rangle} John_2 \{_{INFLP} INFL, \{_{v^*P} John_1, \{_{v^*P} \boxed{what_2}, \{_{v^*P} v^*, ... what_1 ... \}}\}}\}}\}}\}$$

$$(38) \quad WS = \{_{CP} C_{[Q]} \{_{\langle Q, Q \rangle} who_2 \{_{INFLP} INFL, \{_{v^*P} who_1, ... \{_{CP} C\text{-(that), } \{_{\langle \phi, \phi \rangle} who_2 \{_{INFLP} INFL, \{_{v^*P} who_1, ... \}}\}}\}}\}$$

However, these derivations are illegitimate for following reason: Since $what_2$ in (37) and who_2 in (38) remain *in situ* at Externalization and undergo *wh*-movement only at the CI interface, the matrix $C_{[Q]}$ does not have any effect from its access to $what_2$ and who_2 .¹⁵ In contrast, the derivations of (39) and 40) are legitimate because the matrix $C_{[Q]}$ is accessible to the matrix clause *wh*-phrase.

$$(39) \quad WS = \{_{CP} C_{[Q]} \{_{\langle Q, Q \rangle} \boxed{what_2} \{_{INFLP} INFL, \{_{v^*P} who_1, ... \{_{CP} C\text{-(that), } \{_{\langle \phi, \phi \rangle} John_2 \{_{INFLP} INFL, \{_{v^*P} John_1, \{_{v^*P} v^*, ... what_1 ... \}}\}}\}}\}$$

¹⁵ Aoun and Li (1993: 232-233) state that interrogative features comprise three components: i) a Qu-feature, which indicates the interrogative aspect of a clause; ii) a *wh*-feature, which allocates scope to a phrase; and iii) a Focus-feature, which emphasizes particular keywords.

- a. [+Qu, +wh] = *wh*-question b. [+Qu, -wh] = yes/no question
 c. [-Qu, -wh] = statement d. [-Qu, +wh] = exclamatory statements

$$(40) \quad \text{WS} = \{ \text{CP } C_{[Q]} \{ \langle Q, Q \rangle \text{ whO}_2 \{ \text{INFLP INFL}, \{ \text{v*P whO}_1, \dots \} \{ \text{CP } C\text{-}(that), \{ \langle \phi, \phi \rangle \text{ whO}_2 \{ \text{INFLP INFL}, \{ \text{v*P whO}_1, \dots \} \} \} \} \}$$

To recapitulate this section, whether a matrix phase head *C* can access a *wh*- or topic element determines the possibility of extracting the *wh*- or topic element from an embedded clause. If a shared $\langle Q, Q \rangle$ or $\langle \text{Topic}, \text{Topic} \rangle$ label is projected within an embedded clause, the matrix *C* is prevented from accessing the *wh*- or topic element, thereby blocking its extraction from the embedded clause.

4. Conclusion

This paper has argued that the asymmetric behaviors in terms extraction of subject and object operators from an embedded clause arise from the accessibility relation between a matrix phase head *C* and its goal. When an overt complementizer *that* transfers its force feature (subsuming [Q] or [Topic]) to INFL, agreement within the embedded {XP, INFLP} yields a shared $\langle F, F \rangle$ label, making further access by the a matrix phase head *C* unnecessary and thus blocking extraction—a pattern manifested as the *that*-trace effect. Conversely, in the absence of the overt complementizer *that*, a matrix phase head *C* can access its goal, producing grammatical extractions.

References

- Albuhayri, S. 2019. *Information Structure in Standard Arabic Verbal Sentences*. Doctoral Dissertation, University of Wisconsin-Milwaukee.
- Aoun, J. and A. Li. 1993. Wh-Elements in Situ: Syntax or LF?. *Linguistic Inquiry* 24(2), 199-238.
- Chomsky, N. 2013. Problems of projection. *Lingua* 130, 33-49.
- Chomsky, N. 2015. Problems of projection: Extensions. In E. Di Domenico, C. Hamann and S. Matteini, eds., *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, 223, 1-16. Amsterdam: John Benjamins.
- Chomsky, N. 2019a. MIT Lectures. Available online at <http://whamit.mit.edu/2019/05/06/noam-chomskys-lectures-now-online>.
- Chomsky, N. 2019b. Some Puzzling Foundational Issues: the Reading Program. *Catalan Journal of Linguistics Special Issue*, 263-285.
- Chomsky, N. 2020. The UCLA Lectures. Retrieved from <https://ling.auf.net/lingbuzz/005485>.
- Chomsky, N. 2021. Minimalism: Where are we now, and where can we hope to go. *Gengo Kenkyu* 160, 1-41.
- Chomsky, N. 2023a. The Miracle Creed and SMT. [Manuscript]. Available online at <http://www.icl.keio.ac.jp/news/2023/Miracle%20Creed-SMT%20FINAL%20%2831%29%201-23.pdf>
- Chomsky, N. 2023b. Working Toward the Strong Interpretation of SMT. [Online Lecture]. Available online at <https://www.youtube.com/watch?v=n6r957AgzDw>
- Chomsky, N. 2024b. Miracle Creed and SMT. In M. Greco and D. Mocchi, eds., *A Cartesian Dream: A Geometrical Account of Syntax*, 17-39. Lingbuzz Press.
- Chomsky, N., T. D. Seely, R. C. Berwick, S. Fong, M. A. C. Huijbregts, H. Kitahara, A. McInnerney and Y. Sugimoto. 2023. *Merge and the Strong Minimalist Thesis*. Cambridge: Cambridge University Press.
- Emonds, J. 2004. Unspecified categories as the key to root constructions. In A. Adger, C. de Cat & G. Tsoulas,

- eds., *Peripheries: syntactic edges and their effects*. Dordrecht: Kluwer.
- Hisatsugu, K. and D. Seely. 2025. Simplifications and extensions of the Miracle Creed. *Linguistic Variation*.
<https://www.jbe-platform.com/content/journals/10.1075/lv.00040.kit>. <https://doi.org/10.1075/lv.00040.kit>
- Kwon, K.Y. 2025. Form Copy(FC)-Based Approach and Its Consequences. *Journal of Language Sciences* 32(2), 175-194.
- Lasnik, H. and M. Saito. 1993. *Move a: Conditions on Its Application and Output*. Cambridge, Mass.: MIT Press.
- Nakashima, T. 2025. Subject-Object Asymmetries in Box Theory. *Tokai English Studies* 7, 23-41.
- Perlmutter, D. 1971. *Deep and Surface Structure Constraints in Syntax*. New York, NY.: Holt, Rinehart & Winston.
- Pesetsky, D. 1982. *Paths and Categories*. Doctoral dissertation, Massachusetts Institute of Technology.
- Pesetsky, D. 2017. Complementizer-trace effects. In M. Everaert and H. van Riemsdijk, eds., *The Wiley Blackwell Companion to Syntax* (2nd edition), 1-34. New Jersey: John Wiley and Sons Inc.
- Romaine, I. 2015. *A Phase Approach to Spanish Object Clitics*. Doctoral Dissertation, University of California-Los Angeles.
- Seely, T. D. and H. Kitahara. 2025. Simplifications and Extensions of the Miracle Creed. In Andreas Blümel, ed., *Special Issue: Language Variation*. John Benjamins.
- Shim, J. Y. 2019. *Understanding Minimalism in Linguistic Theory*. Bucheon, Korea: Bookk press.
- Sobin, N. 1987. The variable status of Comp-trace phenomena. *Natural Language and Linguistic Theory* 5, 33-60.

Examples in: English

Applicable Languages: English

Applicable Level: All