A Label-based Account for the *That*-trace Effect

Ryota Nakanishi

1. Introduction

One of the most investigated issues in the literature of the generative syntax is *that*-trace effect. Since Perlmutter (1968) reports first, a large number of researchers have been paying much attention to that phenomenon. As the word itself shows, the sequence of an overt complementizer and a trace may cause deviancy.[1] (1) illustrates the effect in English.

(1) Who$_i$ do you think (*that) t_i bought the book?

Here the *wh*-phrase originates as a subject of the embedded clause and undergoes movement to the sentence-initial position. In the presence of the complementizer *that*, the whole sentence is deviant while in its absence, it does not show any grammatical oddness. This ungrammaticality contrasts with the *wh*-movement of an object, where the whole sentence is grammatical regardless of the presence or absence of the complementizer, as shown below:

(2) Who$_i$ do you think (that) John met t_i ?

Thus, *that*-trace effect observes only with the movement of a subject. The same observation can be found in many other language than English, such as French, Russian, Nupe to name a few (see Perlmutter 1971 for French, Pesetsky 1982 for Russian, Kandybowicz 2006 for Nupe)

Numerous syntactic accounts for this effect have been proposed in the literature. For instance, Chomsky and Lasnik (1977) argue that sentences such as (1) are excluded by a filter, as given in (3), which prevents the adjacent occurrence of a complementizer and a trace.

(3) ***That-trace Filter***

*[$_{CP}$ that/whether/if t_{wh} …]

Pesetsky (1982) also presents an analysis, according to which the *that*-trace effect is explained by the Doubly Filled Comp Filter (cf. Chomsky and Lasnik 1977) and the Nominative Island Condition (cf. Chomsky 1980); hence only the case in which a complementizer preceding a subject trace survives,

[1] It is also know that there are cases where such sequence does not become deviant. As shown in (i), subject relativization is perfectly grammatical with an overt complementizer; rather it is obligatory.

(i) the man *(that) met John

However, due to space limitation, we will focus only on cases such as (1) and (2), and let us tentatively leave aside cases such as (i) in this paper (but see Douglas 2015 for the discussion of the anti-*that*-trace effect).

satisfying both the filter and condition.

Although these analyses might be successful, the main purpose of this paper is not to examine them in depth. Rather, we will see how the effect in question is given a theoretical account within the recent framework of the Minimalist Program.

This paper is organized as follows. With a brief introduction of the *that*-trace effect in Section 1, we will examine the analyses by Chomsky (2015) and Erlewine (2017) for the purpose stated just above. After evaluating them, we will then point out the potential problem for both accounts. In Section 3, we will propose that the same configurations represented in terms of labels exclude movement. It will be shown that this proposal can give a coherent account to the problem discussed in the previous section. Section 4 concludes this paper.

2. Previous Studies

In this section, we will first overview Chomsky's (2015) analysis which is based on the Labeling Algorithm (LA) (see also Chomsky 2013) for the explanation of the *that*-trace effect. We will also evaluate Erlewine's (2017) analysis whose argument is extended from Erlewine (2014, 2016).

2.1. Chomsky (2015): LA-based Account

Chomsky (2015) claims that the *that*-trace effect is explained as a consequence of the theory of Labeling Algorithm (LA). Given the assumption that every syntactic object must be regarded to be legitimate, he argues that it must be assigned a label by LA at syntax; to be more precise, LA is a minimal search which inspects the internal structure of a syntactic object according to which it assigns a label if successful. According to him, an external argument must raise to Spec TP; otherwise the labeling of *v**P fails and it cannot be properly interpreted at the interfaces. Thus, the raising of an external argument makes it possible for TP to be assigned a label <φ, φ> and for *v**P to be labeled as *v**, deriving the traditional EPP effect. Of importance here is that TP can be labeled because of the presence of the raised external argument, and that *v**P also poses no labeling failure since the external argument is raised and hence it is not targeted by LA. Chomsky then assumes that the same derivation is involved when an external argument is a *wh*-phrase: a *wh*-external argument must be raised to Spec TP. For the purpose of illustrating the (un)grammaticality of (1), now suppose that we have reached a derivational stage where the *wh*-external argument *who* is raised to Spec TP because of the successful labeling and C is then merged, as shown below:

(4) [C [$_{\alpha}$ who$_i$ T [$_{\beta}$ t_i *v** […]]]]

As stated just above, the labeling of the syntactic object β is a successful since *who* is raised and the

trance of it is invisible for LA. Assuming that *who* also has a φ-feature, here too, α successfully gets a label <φ, φ>. However, he argues that the derivation after this stage will crash if C is an overt complementizer such as *that*.

Now suppose that C is overt. Chomsky also assumes that the complement of a phase head is the domain of the operation Transfer and that all the elements inside it are transferred to the interfaces and become inaccessible for further syntactic operations (cf. the Phase Impenetrability Condition; PIC). In the case of (5), the relevant domain is α (the shaded part) because C is a phase head, so that if nothing more happens, *who* would be transferred along with the other elements inside α, resulting in a derivational crash.

(5) [that(C) [$_{\alpha}$ who$_i$ T [$_{\beta}$ t_i *v** […]]]]

Here the difference between the overt and null complementizer comes into play. Suppose now that C is null. Chomsky assumes that the null complementizer obtains by deleting C, so that C disappears from the syntactic computation. Since now that C does not exist, the natural assumption is that the phasehood that C had is inherited to T through Feature Inheritance (of φ-feature) and activates on T. Thus T acts a phase head and the domain of Transfer slides from α to β. The structure after the C deletion is given in (6) (Transfer applies to the shaded part).

(6) [$_{\alpha}$ who$_i$ T [$_{\beta}$ t_i *v** […]]]

As *who* is outside of the domain, it can remain in Spec TP and is still available for further syntactic operations, undergoing successive cyclic movement. Therefore, only when C gets deleted can a *wh*-external argument undergo further movement, not observing the *that*-trace effect.

To summarize, a *wh*-external argument is allowed to undergo further movement in the absence of a complementizer because it can escape the Transfer domain by means of the inheritance of phasehood, whereas it cannot escape in the presence of an overt complementizer because it is inside the Transfer domain.

2.2. Erlewine (2017): Anti-locality-based Account

Before we evaluate Chomsky's (2015) analysis, let us examine Erlewine (2017) which offers an explanation from a different view point.

Based on the extraction asymmetry in Kaqchikel in Erlewine (2014, 2016), Erlewine (2017) claims that there is a locality constraint which bans too short movement, which he calls Spec-to Spec Anti-locality , as stated in (7a). He also defines the "crossing" in this constraint as in (7b).

(7) a. ***Spec-to-Spec Anti-locality***

A'-movement of a phrase from the Specifier of XP must cross a maximal projection

other than XP.

b. ***Definition: crossing***

Movement from position α to position β *crosses* γ if and only if γ dominates α but does not dominate β.

(Erlewine 2017:373)

This constraint states that movement from a specifier position to a higher specifier position immediately above it is too short because there is no intervening maximal projection and thus is prohibited. Armed with these, the ungrammaticality of (1) straightforwardly follows. Consider the following derivation:

(8) $[_{CP}$ who$_i$ that $[_{TP}$ t_i T […]]]

In (8), after the wh-external argument *who* is raised to Spec TP, it is raised further to Spec CP to undergo successive cyclic movement. However, since this movement is from Spec TP to Spec CP, it is subject to the constraint in (7a), resulting in the ungrammaticality; the *that*-trace effect observes.

This analysis is supported by the adverb insertion effect. It has been reported that in some languages such as English, Nupe, and Yiddish, the that-trace effect is obviated by inserting an adverb between a complementizer and a trace (see see Kandybowicz 2006 for Nupe, Erlewine 2017 for Yiddish). English shows this effect, which is exemplified in (9)

(9) a. *Who$_i$ did John say that t_i ran to the store?

b. Who$_i$ did John say that fortunately t_i ran to the store? (Erlewine 2017:375)

Assuming that an adverb has a projection of its own, (9b) has the following structure:

(10) $[_{CP}$ who$_i$ did John say $[_{CP}$ t_i that $[_{AdvP}$ fortunately $[_{TP}$ t_i ran to the store]]]]?

The movement from Spec TP to Spec CP in (10) crosses an intervening maximal projection, i.e. AdvP, so that it is long enough to survive the constraint in (7a) in contrast to (8).

What is important here is that movement from Spec TP to Spec CP is banned regardless of whether C is overt or null. Here, a natural question arises: why is the sentence with a null complementizer not subject to the locality constraint? Following Fox and Pesetsky (2005), Erlewine argues that the key lies in Cyclic Linearization. Fox and Pesetsky suggest that linearization is determined cyclically within each Spell-Out domain and that an ordering paradox crashes a whole derivation. For instance, once an ordering relation X < Y (where "<" means "precedes") is established, no contradicting ordering relation Y < X should be obtained; otherwise, we get an ordering paradox. What he emphasizes with this mechanism is that it is possible for a *wh*-external argument to undergo direct movement from an embedded clause to a matrix clause, without dropping at an intermediate Spec CP. For the illustration of how this works, consider the following derivation of (1) with a null

complementizer, where *who* directly moves from the embedded TP to the matrix Spec CP:

(11) a. [$_{CP}$ who$_i$ do you think [$_{CP}$ Ø [$_{TP}$ t_i bought the book]]]

b. *Linear order relations in embedded CP:* who < bought < the < book

c. *Linear order relations in matrix CP:* who < do < you < think < CP

⇒ no ordering paradoxes

Because the complementizer is phonetically null, it does not join the determination of linearization and we get the ordering relations in (11b) at the embedded CP level. After *who* moves to the matrix Spec CP, ordering relations are calculated again, obtaining those in (11c) at the matrix CP level. Since the ordering relations in (11b) and (11c) does not contradict, the linearization succeeds. Notice that on the other hand, such a direct movement in the presence of an overt complementizer is not allowed for linearization, as shown in (12).

(12) a. [$_{CP}$ who$_i$ do you think [$_{CP}$ that [$_{TP}$ t_i bought the book]]]

b. *Linear order relations in embedded CP:* that < who < bought < the < book

c. *Linear order relations in matrix CP:* who < do < you < think < CP

⇒ ordering paradox! (who < that *vs.* that < who)

Unlike (11), the complementizer is an element which has a phonetic content, so that it is subject to linearization. As a result, as (12c) indicates, we have the contradicting order relations: *who* cannot precede and follow *that* simultaneously. This is why a null complementizer allows the direct movement of a *wh*-external argument from Spec TP while an overt counterpart does not.

To recapitulate, according to Erlewine's analysis, the two possible derivations with an overt complementizer are both excluded: the movement through an intermediate Spec CP is prevented by the Spec-to-Spec Anti-locality and the direct movement from Spec TP is also ruled out by Cyclic Linearization.

2.3. Potential Problem

In the previous subsections, we have overviewed the analyses by Chomsky (2015) and Erlewine (2017). Although their accounts have different analytic view points from each other, a common expectation follows when we examine them, which will be discussed carefully in this section.

Their accounts share one point, namely that in the absence of an overt complementizer, a *wh*-external argument undergoes movement directly from an embedded Spec TP to a matrix clause. This raises a natural prediction that there is no intermediate landing site for a *wh*-external argument. Thus the movement they both assume is not strictly cyclic as has standardly assumed in the literature; it must skip an intermediate Spec CP to avoid unwelcome results. However, it has widely been argued

that there is evidence that successive cyclic movement should go through such a position. Observe the following examples, especially the successful coreference in (13b):

(13) a. *Mary told John$_i$ that she would buy these pictures of himself$_i$.

b. [Which pictures of himself$_i$]$_j$ did Mary tell John$_i$ that she would buy t_j ?

(13a) shows that the coreference between *John* and *himself* is impossible because of a violation of Condition A. However, once the reflexive is moved as in (13b), the ungrammaticality disappears and the intended coreference becomes possible. Since the antecedent needs to c-command the reflexive for successful binding, the reflexive should drop at some position where it can get c-commanded by the antecedent. Given that the surface position does not satisfy this requirement, there should be other position than the surface and base position. Therefore, we are led to conclude that *wh*-movement must go through the embedded Spec CP where the reflexive can get c-commanded by the antecedent.

The natural prediction raised just above can now be checked with this line of argument. If the movement of a wh-external argument does not go via an intermediate Spec CP, it is predicted that the reflexive embedded in an embedded *wh*-external argument cannot be coreferential with an antecedent in a matrix clause. However, this is not borne out.

(14) a. *Mary told John$_i$ (that) these pictures of himself$_i$ surprised Bill.

b. ?[Which pictures of himself$_i$]$_j$ did Mary tell John$_i$ t_j surprised Bill?

(p.c. Connor Mayer, Jeremy Steffman)

In (14), the intended reading is possible, which indicates that the *wh*-phrase should go through the embedded Spec CP, where the reflexive can be bound by its antecedent. Therefore, even in the absence of a complementizer, Spec CP must be filled in the course of derivation regardless of whatever undergoes long-distance *wh*-movement, contra Chomsky (2015) and Erlewine (2017).

Summing up so far, it has been shown in this section that although their analyses can capture the effect, they make a wrong prediction with respect to possible binding relation. For this reason, we are led to seek an alternative account for the *that*-trace effect and the successful binding discussed above.

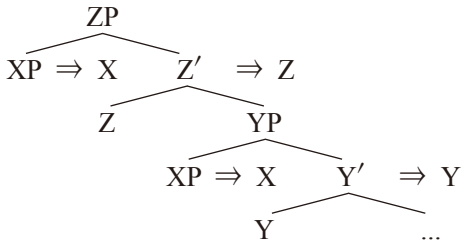
3. Proposal

Seeking an alternative account, this section develops and proposes an analysis which is based on labels. As briefly introduced in Section 2.1., according to Chomsky, every syntactic object is assigned a label so that it is appropriately interpreted at the interfaces. If this assumption is correct, it is reasonable to assume that labels play an important role to check whether a syntactic object mapped from syntax is legitimate. Thus, in this paper, we take a label-based approach to explain the empirical

data discussed in the previous section.

To be more precise, I propose that if a configuration that movement creates results in the same configuration in terms of labels as the one before it applies, such movement is not allowed. For the illustration, consider the following configuration where XP moves from Spec YP to Spec ZP:

(15)



Here, Z and Y are the heads of ZP and YP, respectively, hence Z′ and Y′ being labeled as Z and Y, respectively. Suppose now that XP's label is its head, X. Then, following Chomsky's (2004 *et seqq*) idea that Merge creates a set, ZP and YP are set-theoretically represented in terms of labels as {X, Z} and {X, Y}, respectively. Thus, before movement, XP resides in the configuration {X, Y} but after movement, XP does in {X, Z}. We also assume that movement creates a chain, which is represented with an ordered set consisting of configurations the moved element resides in. The movement chain in (15) is thus represented as <{X, Z}, {X, Y}>. With these said, now our proposal that if the configurations before and after movement is the same, such movement is not allowed is formalized as follows:

(16) *<{X, Y}, {X, Z}> iff Y = Z

The leading idea utilized in our proposal is that such exclusion follows from the Principle of Economy, which we take here to mean that if movement does not create a new configuration, i.e. results in the same configuration, then such movement is so trivial that economical consideration excludes it. Put differently, we are saying that no trivial movement must be included. Therefore, identifying Y with the same label as Z leads to ungrammaticality.

In addition to this proposal, we assume the mechanism of Feature Inheritance in the sense of Chomsky (2008, 2013, 2015), as mentioned in Section 2.1. Although Chomsky presumes that after the Feature Inheritance, φ-features are no longer present on C, we assume that they are. In other words, through the Feature Inheritance, the φ-features on C are copied into T and they are present on both C and T. We assume further that once the φ-features on T are valued by agreement between T and a subject, its valuation is transmitted to C through the relation established by the Feature Inheritance, thereby the ones on C are also valued.

Armed with these proposal and assumptions, now let us see how they work to achieve our goal. Suppose we have reached the CP phase level, where the φ-features on C are copied onto T by the

Feature Inheritance and the *wh*-subject *who* moves from Spec TP to Spec CP, as illustrated in the structure in (17).

(17) $[_{CP}\ who_{[\varphi]}\ [_{C'}\ \ that_{[\varphi]}(C)\ [_{TP}\ <who_{[\varphi]}>\ [_{T'}\ \ T_{[\varphi]}\ [\ \ldots\ ]]]]]$

The *wh*-subject *who* in Spec TP agrees with T and the prominent feature shared by them, namely the φ-feature, becomes the label of TP and subsequently T′ is also assigned the same label.[2] Moreover, by our assumptions, C has the φ-features that have the same values as the ones on T. Thus the minimal search determines the CP's label as φ, the prominent feature share by *who* in Spec CP and C, and, accordingly, C′ is assigned the label φ. Now the structure in (17) is identical to that in (15) with respect to labels. Before movement, the configuration in which *who* in Spec TP resides is {φ, φ} but after movement, the configuration is the same as before. Therefore, the movement in (17) is excluded and we observe the *that*-trace effect as in (1). Notice at this point that we can now exclude the configuration that Erlewine's analysis also does. In other words, our proposal can exhibit the same empirical coverage as his analysis.

Crucially, our current proposal hinges on the presence of φ-features on C, which make C identical to T with respect to their labels. This predicts that if φ-features on C disappear, the movement from Spec TP to Spec CP is allowed since the configuration after movement is no longer identical to the one before it. We then claim that this is what we have observed in (1), where deleting *that* salvages the sentence. Here we take the operation that deletes *that* to be an operation that deletes a phonetically related content. This is motivated the fact that φ-features are related to a phonetic content because, for example, verb forms depend on the φ-features subjects have, e.g. third person singular for *kicks*, first person plural for *kick*, etc. Thus we suppose that the deletion of *that* also gets rid of φ-features. Assuming that after the deletion of *that*, C, not φ, becomes the label of both CP and C′, then the deletion avoids resulting the same configuration: before movement, the configuration is {φ, φ} but after that, it is {φ, C}. Hence, no *that*-trace effect.[3]

This line of explanation does not affect the grammaticality of (2). Although moving the *wh*-object to the embedded Spec CP seemingly creates the same situation as in (17), it is, in fact, not the same. Since the φ-features of the *wh*-objects, represented as φ_{obj}, are different from the ones of the

[2] Although Chomsky (2013, 2015) does not explicitly discuss the labeling of intermediate projections such as T′, the assumption that every syntactic object must be successfully labeled requires them to assigned labels. Since he assumes that after agreement, the English T is strengthened so that it can serves as a label, then we assume that T′ is labeled as φ.

[3] In this case, contrary to the one where an overt complementizer is present, there is no feature sharing for labeling. However, this does not cause a labeling failure for the embedded CP. Given that traces are invisible for LA (cf. Chomsky 2013, 2015), since after that level the *wh*-subject moves from the embedded Spec CP, it becomes possible for LA to successfully determine CP's label as C.

subjects, φ_{subj}, the movement in question does not result in the identical configuration: given that the *wh*-object agrees with V and then moves to the embedded Spec CP, the configuration before movement is $\{\varphi_{obj}, \varphi_{obj}\}$ but the one after it is $\{\varphi_{obj}, \varphi_{subj}\}$. Therefore this movement is not ruled out by (16). The structure is given in (18), where only the relevant parts are illustrated.

(18) $[_{CP}$ who$_{[\varphi obj]}$ $[_{C'}$ that$_{[\varphi subj]}$(C) $[_{TP}$ John$_{[\varphi subj]}$ $[_{T'}$ T$_{[\varphi subj]}$ $[_{vP}$ $v_{[\varphi obj]}$ $[_{VP}$ <who$_{[\varphi obj]}$> $[_{V'}$ V$_{[\varphi obj]}]]]]]]]$

Note that our proposal does not rule out movement from Spec TP to Spec CP *per se*, which Erlewine excludes with the Anti-locality constraint. It is only when there is an overt complementizer that such movement is excluded. This point explains the availability of the interpretation we have observed in (14b). Since, in our proposal, nothing prevents the *wh*-subject containing the reflexive from moving to the embedded Spec CP in the absence of *that*, then the reflexive can be bound by its antecedent and interpreted there. Hence the intended interpretation straightforwardly follows. This cannot be achieved by Chomsky's and Erlewine's analyses where the *wh*-subject moves directly from the embedded Spec TP, skipping the embedded Spec CP. Considering the availability of that interpretation as well as the same empirical coverage as Erlewine's analysis, it can be concluded that our analysis is one step ahead.

4. Conclusion

Throughout this paper, the *that*-trace effect and its analysis have been examined from the view point of the recent Minimalist Program. In Section 2, we have overviewed and evaluated the previous analyses; in particular, Chomsky (2015) and Erlewine (2017). Their accounts appeared to capture the effect well and succeed in giving theoretical explanations to it. However, it has also shown that their prediction did not hold and their accounts have been proved to be insufficient regarding the binding possibility. In this context, Section 3 have offered an analysis which makes use of labels and according to which resulting in the same configuration as before leads to ungrammaticality, being excluded by economical consideration. Also, our analysis captures the successful binding relation which cannot be predicted by Chomsky's and Erlewine's analyses: the embedded CP becomes available as an intermediate position of movement when φ-features on C along with its phonetic content are deleted.

References

Chomsky, Noam (1980) On binding. *Linguistic Inquiry* 11:1–46.

Chomsky, Noam (2008) On phases. In: Robert Freidin, Carlos P. Otero and María Luisa Zubizarreta (eds.) *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–

166. Cambridge, MA: MIT Press.

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

Chomsky, Noam (2015) Problems of projection: Extensions. In: Elisa Di Domenico, Cornelia Hamann and Simona Matteini (eds.) *Structures, strategies and beyond: Studies in honor of Adriana Belletti*,1–16. Amsterdam: John Benjamins Publishing Company.

Chomsky, Noam and Howard Lasnik (1979) Filters and control. *Linguistic Inquiry* 8:425–504.

Douglas, Jamie A. (2015) Unifying the *that*-trace and anti-*that*-trace effects. ms., University of Cambridge.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely (2016) Phase cancellation by external pair-merge of heads. *The Linguistics Review* 33: 87–102.

Erlewine, Michael Yoshitaka (2014) Anti-locality and Kaqchikel agent focus. In: Robert E. Santana-LaBarge (ed.) *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 150–159. Somerville, MA: Cascadilla Proceedings Project.

Erlewine, Micheal Yoshitaka (2016) Anti-locality and optimality in Kaqchikel agent focus. *Natural Language and Linguistic Theory* 34: 429-479.

Erlewine, Michael Yoshitaka (2017) Why the null complementizer is special in complementizer-trace effect. In: Claire Halpert, Hadas Kotek and Coppe can Urk (eds.) *A pesky set: Papers for David Pesetsky*, 371–380. Cambridge, MA: MIT Working Papers in Linguistics.

Fox, Danny and David Pesetsky (2005) Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–45.

Kandybowicz, Jason (2006) *Comp-trace* effects explained away. In: Donald Baumer, David Montero and Michael Scanlon (eds.) *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 220–228. Somerville, MA: Cascadilla Proceedings Project.

McCloskey, James (1996) Resumption, successive cyclicity, and the locality of operations. In: Samuel D. Epstein and T. Daniel Seely (eds.) *Derivation and explanation in the Minimalist Program*, 184–226. Oxford: Blackwell.

Perlmutter, David M (1968) Deep and surface structure constraints in syntax. Doctoral dissertations, MIT.

Perlmutter, David M (1971) *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Pesetsky, David (1982) Complementizer-trace phenomena and the nominative islands condition. *The Linguistic Review* 1: 297–344.