

Subject Extraction and Clause Size*

MANABU MIZUGUCHI

Toyo University

mizuguchi@toyo.jp

Abstract: It has been noted in the literature that subject extraction shows peculiarities. This paper discusses the argument that when subject extraction takes place, the embedded clause out of which the subject is moved will always be a smaller-than-full clause, and addresses why such a clause matters for subject extraction. Assuming the Strong Minimalist Thesis (SMT) to be the basic hypothesis in quest of an explanatory theory of language, I claim that subject extraction can be explained by the structure building operation Merge and the labeling algorithm called Label, both of which are elements of principled explanation that follow from SMT, answering why clause size counts for extraction of the subject. I argue that the proposed analysis provides the simplest, hence principled, account of language. I also show that it brings three consequences, which endorse the role of Merge in language.

1. Introduction

Pesetsky (2016) argues for the following generalization as regards subject extraction:

- (1) Subject extraction always entails a smaller-than-full clause.

According to this generalization, in (2a,b) and (3a), the embedded clause is less-than full:

- (2) a. Which student do you believe [to be the most intelligent]?
b. Which student do you believe [is the most intelligent]?
c. * Which student do you believe [that is the most intelligent]?
- (3) a. The student seems [to be in the library].
b. * The student seems [that is in the library].

Pesetsky proposes Exfoliation, which deletes and removes a certain portion of the structure to explain the generalization. He argues that in the relevant examples, unlike in (2c) and (3b), CP is removed by Exfoliation and that the subject is extracted out of a smaller-than-full clause:

* I would like to thank the audience at FLYM 2018 for comments and questions on this paper. All remaining errors and inadequacies are my own. The research reported in this paper has been conducted under Grant-in-Aid for Young Scientists (B) (#16K16835) from Japan Society for the Promotion of Science (JSPS).

- (4) ... [VP V [CP C [TP T [VP v [...]]]]]
- ↑
- Exfoliated

In fact, the generalization goes back to Chomsky (1981), where the rule of S'-deletion or more generally, the S'-to-S rule, is proposed to account for subject extraction in sentences such as (5):

- (5) a. Who do you think [_s saw Bill]?
- b. Bill was believed [_s to have seen Tom].

More recently, Chomsky (2015) proposes *that*-deletion, which deletes C and removes a C-headed set or CP, to account for the contrast between (2b) and (2c).¹

The purpose of this paper is to consider the generalization (1), addressing why clause size matters for subject extraction and how clause reduction for subject extraction is executed. I argue that the subject-extraction generalization is explained by Merge and Label, claiming that it receives a principled explanation under the Strong Minimalist Thesis (SMT).

2. Theoretical Background

In this section, I first discuss a framework of principled explanation and introduce theoretical background I assume in this paper. The Basic Property of the Faculty of Language (FL) is that a language yields a digitally infinite array of hierarchically structured expressions or syntactic objects (SOs) with systematic interpretations at interfaces with the conceptual-intentional (CI) system and the sensorimotor (SM) system (Chomsky 2017:295). Given that linguistics is a branch of science and that FL must meet the condition of evolvability, which implies that Universal Grammar (UG) is quite simple at its core, the basic hypothesis in an explanatory theory of FL is that the simplest account should be sought. Given the Basic Property, the best scenario would be that phenomena are explained only by an operation that generates hierarchically structured SOs, which is called Merge in the literature, and by interfaces with the CI and SM systems, which subject derived SOs to semantic and phonological interpretations. This theoretical hypothesis is articulated by the Strong Minimalist Thesis (SMT) (=6) (Chomsky 2000 et seq., especially Chomsky 2010, 2017):

- (6) Language = Merge + interfaces

Merge interacts with general principles of minimal computation, which constrain computation in general. Merge, constrained only by such principles, falls out as simplest Merge and applies freely to any two SOs to form a new SO out of them. At the interfaces, Full Interpretation must be satisfied; otherwise, interface conditions will be violated. The bare minimum requirement for Full Interpretation is identification of SOs created by Merge, which are label-less, and hence unidentified since Merge simply puts two SOs together. CI and externalization at the SM level must know what

¹ See also Bošković (1997) and Ishii (2004), where it is argued that the *that*-less clause in (2b) is TP.

kind of object a derived SO is and SOs must be identified for interpretation (identifiable = interpretable at the interfaces). Chomsky (2013, 2015) proposes that identification of SOs is carried out by a labeling algorithm called Label, which is simply an instantiation of minimal search. Applied to an SO, Label locates the closest head (in the case of a set of the form X-YP) or agreeing heads (in the case of an XP-YP set), both of which are equally close to Label, determining its properties or label and providing necessary information about it for interpretation at CI and for externalization. Labeling by Label is one necessary component of Full Interpretation and is motivated by the interfaces. As far as language is explained by Merge and Label, the simplest, hence principled, account of language will be achieved, and language will be a perfectly designed system.

With this background in place, let us now consider why a smaller-than-full clause matters for subject extraction.

3. Subject Extraction from a Full Clause

I first discuss extraction from a full clause, considering why the derivation of subject extraction will be ill-formed if the embedded clause is fully structured. Take (2c) for illustration, keeping in mind that the same argument applies to (3b). In the derivation of SOs, I assume a phase-based model of derivation proposed by Chomsky (2007, 2008), in which all operations except External Merge (EM) apply at the phase level. Given this assumption, extraction or internal merge (IM) of the subject will apply when the C phase is structured in the embedded clause. Consider (7):

(7) $[\kappa \text{ that } [\lambda \text{ T } [\alpha \text{ be } [\text{which student } [\text{the most intelligent}]]]]]$

At (7), two derivational options are available for subject extraction. In Derivation 1, the subject moves to Spec,TP. This option, however, will halt subject extraction in the embedded clause. Though details of their proposals differ, Epstein, Kitahara and Seely (2012) and Mizuguchi (2014b) both claim that movement to Spec,TP at the phase level will induce Transfer as it applies counter-cyclically, arguing that Transfer allows the movement to apply without causing a derivational failure. Details aside, they argue that movement to Spec,TP, which counter-cyclically merges the subject with λ , will transfer γ (= (8a)). If γ is transferred, the subject will be transferred and will get inaccessible to computation in the higher clause since transferred domains are syntactically forgotten, with the result that syntax cannot move it out of the embedded clause (= (8b)):²

(8) a. $[\kappa \text{ that } [\gamma \text{ which student } [\lambda \text{ T } [\alpha \text{ be } [\text{t } [\text{the most intelligent}]]]]]]]$
 \Rightarrow Transfer

b. $[\text{---} [\kappa \text{ that } [\gamma \text{ which student } [\lambda \text{ T } [\alpha \text{ be } [\text{t } [\text{the most intelligent}]]]]]]]$
 \uparrow X INVISIBLE

In Derivation 1, subject extraction will be derivationally ruled out.

Now consider the other derivational option. In Derivation 2, the subject moves to Spec,CP

² “t” in (8) and elsewhere in this paper is conveniently used as a copy of a moved SO.

directly, which allows it to avoid being transferred as the movement is cyclic and applies to the root; σ , unlike γ in (8b), is not transferred upon the movement of the subject:

- (9) $[\sigma \text{ which student } [\kappa \text{ that } [\lambda \text{ T } [\alpha \text{ be } [\text{t } [\text{the most intelligent}]]]]]]]$

Under this derivation, however, λ will not be labeled and labeling failure will result. In English, T alone is too weak to serve as a label and cannot label on its own. In order to work as a label, it must have overt or visible Spec,TP when Label applies (Chomsky 2015:9-10). In (9), there is no overt element in Spec,TP and λ will not be labeled, even though T can be located by Label through minimal search. Under Derivation 2, an unlabeled SO will be generated and subject extraction will be ruled out at the interfaces as λ cannot be interpreted, violating Full Interpretation.

In summary, the discussion so far has demonstrated that the ill-formedness of subject extraction from a full clause is explained by Merge and Label, the elements of principled explanation.

Notice that simultaneous, parallel movement cannot explain subject extraction in (2) and (3). Chomsky (2008) argues that the subject moves to Spec,CP and Spec,TP simultaneously at the phase level, creating two independent chains and two occurrences of the subject. Consider (10):

- (10) $[\tau \text{ subject } [\kappa \text{ C } [\gamma \text{ subject } [\lambda \text{ T } [\alpha \text{ t } \dots]]]]]]$
-

Suppose that the occurrence created in Spec,TP is a copy. Given that copies are invisible to operations (Chomsky 2013), Label cannot detect overt Spec,TP, with γ and λ left unlabeled. This explains ill-formed cases of subject extraction like (2c) and (3b) but cannot account for well-formed ones such as (2a,b) and (3a). Now suppose that the occurrence in question is not a copy, being visible to Label. If so, γ and λ will be labeled thanks to overt Spec,TP. This explains well-formed cases of subject extraction but not ill-formed ones. In brief, simultaneous movement rules in or out subject extraction across the board and it cannot be a solution to subject extraction.

4. Subject Extraction out of a Smaller-than-Full Clause

4.1. Not by Deletion but by Merge

If subject extraction requires a smaller-than-full clause, then the next questions are: (i) how is the clause reduced and (ii) how does clause reduction make subject extraction possible? As for (i), recall from section 1 that Pesetsky proposes Exfoliation and that Chomsky proposes *that*-deletion. In this section, I propose that a smaller-than-full clause is created by Merge and argue that when the clause is reduced by Merge, the subject can be extracted without inducing transfer and labeling problems; Merge provides a solution to subject extraction.

Recall that Merge is constrained only by general principles of minimal computation (or so-called “third factor” principles) and operates freely. Given that Merge applies freely, two modes of Merge are available for free: that is, set-Merge, which symmetrically merges two SOs and produces a single set (=11a)) and pair-Merge, which asymmetrically merges two SOs and adjoins one SO to the other, yielding an ordered pair (=11b)):

Subject Extraction and Clause Size

- (11) a. $\{\alpha, \beta\}$ b. $\langle \alpha, \beta \rangle$

In (2a,b) and (3a), where the subject can be extracted, I claim that C and T are not set-merged as in (12); instead, as illustrated in (13), T is externally pair-merged to C, which produces an amalgamated head $\langle C, T \rangle$, and then the pair-merged SO is set-merged with α , a set headed by v/v^* .^{3,4}

- (12) a. $[\lambda T [\alpha \dots]]$ b. $[\kappa C [\lambda T [\alpha \dots]]]$

- (13) a. $\langle C, T \rangle$ b. $[\delta \langle C, T \rangle [\alpha \dots]]$

Taking (2a,b) for illustration, under the proposed analysis, (14), instead of (7), is yielded in the embedded clause, where TP is reduced by pair-Merge:

- (14) $[\delta \langle C, T \rangle [\alpha \text{ be } [\text{which student } [\text{the most intelligent}]]]]$

In order to move out into the higher clause, the subject will be internally merged with the δ -marked set, moving to the Spec of $\langle C, T \rangle$; otherwise, it will be trapped in α , which is transferred as a complement of $\langle C, T \rangle$, and cannot move out.⁵ Consider (15a). Notice that the subject will not be transferred upon its movement to the Spec of $\langle C, T \rangle$ since subject movement applies to the root and μ is not transferred for the movement. Moreover, the movement does not cause a labeling problem: in (15a), λ or the T-headed set is not formed and the labeling problem does not arise. The problems we have noted in the last section are solved. The derivation continues to produce (15b), where the subject is extracted. SOs, when pair-merged to others, get de-activated and become syntactically invisible as adjoined SOs are asymmetrical to their hosts (Chomsky 2004). Syntactically, $\langle C, T \rangle$ is on par with C. Since C can label on its own, $\langle C, T \rangle$ can work as a label without overt Spec. In (15b), μ and δ can be labeled when the subject moves out to be a copy:

- (15) a. $[\mu \text{ which student } [\delta \langle C, T \rangle [\alpha \text{ be } [t [\text{the most intelligent}]]]]]$
b. $[\text{which student } [\text{do you believe } [\mu t [\delta \langle C, T \rangle [\alpha \text{ be } [t [\text{the most intelligent}]]]]]]]$

I have demonstrated that clause reduction by pair-Merge allows the derivation of subject extraction to succeed both derivationally and at the interfaces, answering why a smaller-than-full clause matters for subject extraction: thanks to pair-Merge, transfer and labeling problems can be circumvented that would arise under the derivation where C and T are set-merged (=12)). Given that Merge applies freely, the final question is why clause reduction by pair-Merge does not apply

³ Given that Merge applies freely, pair-Merge, like set-Merge, can apply both externally and internally. See Epstein, Kitahara and Seely (2016) for relevant discussion. Unless otherwise mentioned, (pair-)Merge is external (pair-)Merge.

⁴ Erlewine (2017b) proposes that C and T are bundled as CT and argues that this explains a *that-t* violation as it solves anti-locality. See also Erlewine (2017a) and references cited therein for independent arguments for bundled CT.

⁵ As I discuss immediately below, $\langle C, T \rangle$ is syntactically on par with C and hence it works as a phase head.

to (2c) and (3b); in other words, why are C and T set-merged in the relevant examples? I claim that how C and T are merged leads to different realizations of complementizers, arguing that complementizers indicate how C is merged: the English overt complementizer *that* is a spell-out of C, indicating that clause reduction by pair-Merge does not take place; on the other hand, a null complementizer \emptyset and infinitival *to* are externalizations of $\langle C, T \rangle$, with C and T pair-merged in the derivation. Externalization explains why pair-Merge does not occur in (2c) and (3b).

Merge + externalization can also account for subject extraction in Bantu A-movement. Consider (16) from Lubukusu and (17) from Zulu:⁶

- (16) a. Babaandu ba-lolekhana [mbo t ba-kwa].
 2people 2SA-seem that 2SA.PST-fall
'The people seem like they fell/The people seem to have fallen.'
- b. * Mikaeli a-lolekhana [a-li t a-si-kona].
 Michael 1SA-seem that 1SA-PERS-sleep
'Michael seems to be still sleeping.' (Carstens and Diercks 2013)
- (17) a. uZinhle u-bonakala [ukuthi t u-zo-xova ujeqe].
 AUG.1Zinhle 1SA-seems that 1SA-FUT-make AUG.1steamed.bread
'It seems that Zinhle will make steamed bread.'
- b. * uZinhle u-bonakala [t uku-(zo-)xova ujeqe].
 AUG.1Zinhle 1SA-seem INF-(FUT-)make AUG.1steamed.bread
'It seems that Zinhle will make bread.' (Halpert 2015)

In Lubukusu, unlike in English, subject extraction out of the finite clause is possible when a certain overt complementizer is selected; in Zulu, subject extraction out of the non-finite clause is impossible. I argue that the externalization component of Lubukusu and Zulu spells out $\langle C, T \rangle$ as *mbo* / *ukuthi* and C as *a-li* / \emptyset ; in (16a) and (17a), the clause is reduced by pair-merge of T to C. Notice that finiteness is irrelevant to the application of Merge, hence pair-Merge, and that externalization at the sensorimotor level, which is peripheral to UG, is subject to variation within and across languages, inducing apparent complexity and variety of language (Chomsky 2017 among others). The proposal here can explain variation with subject extraction under uniform syntax.

As regards subject extraction in the Bantu examples above, Carstens and Diercks (2013) and Halpert (2015) argue that it is an instance of A-movement. In these examples as well as in (3a), the derivation, like the derivation of (2a,b), also proceeds as illustrated in (18) for successful subject extraction and the subject moves to the Spec of $\langle C, T \rangle$; otherwise, it will be inaccessible to computation as α is cyclically transferred as a complement of $\langle C, T \rangle$, which works as a phase head as it retains the properties of C:

⁶ Abbreviations: SA = subject agreement; PST = past; FUT = future; INF = infinitive; PERS = persistent; AUG = augment. Arabic numerals preceding nouns indicate noun class. *A-li* is an agreeing complementizer and *-li* is the complementizer; *-li* agrees with the most local subject in the higher clause. For details, see Diercks (2013).

Subject Extraction and Clause Size

(18) [C [subject [T [... [_μ t [_δ <C, T> [_α t ...]]]]]]]]

The movement, however, could pose a problem: if the Spec of <C, T> is an A'-position, the subject will undergo improper movement. It has been argued in the literature that such movement is banned, which can be attributed to the fact that it forms a non-uniform chain, causing interpretive ill-formedness at the CI interface (Chomsky and Lasnik 1993, Fukui 1993). The derivation (18) will not be problematic if A/A'-positions are defined by ϕ -features. Following Chomsky (2007, 2008) and Obata and Epstein (2011), I argue that ϕ -features play a key role in determining A-positions, proposing (19) as the definition of such positions:

(19) The *nP* is in an A-position if it is merged with an SO headed by a head bearing ϕ -features; otherwise, it is in an A'-position.

Given (19), the subject moves to an A-position in the embedded clause: it is merged with a set headed by <C, T>, which bears ϕ for under-inheritance. Movement onto the matrix Spec,TP from the Spec of <C, T> will not be improper movement and the derivation will be ruled in.

4.2. Prediction

I have discussed why subject extraction requires a smaller-than-full clause and how clause reduction for subject extraction is executed. It should be noted that the proposed analysis predicts that the subject can be extracted out of a full clause if the transfer and labeling problems can be bypassed. This prediction is borne out. To see this, consider (20) through (22):

(20) Ver hot er moyre [az [*(es) [vet t kumen]]]? Yiddish
 who has he fear that Expletive will come
'Who does he fear will come?' (Diesing 1990)

(21) Vilken elev trodde ingen [att [*(han) [skulle fuska]]]? Swedish
 which pupil thought nobody that he would cheat
'Which pupil didn't anyone think would cheat?' (Engdahl 1982)

(22) Quelles filles crois-tu [qui/*que [vont t acheter ce livre-la]]? French
 which girls think-you that will buy that book-there
'Which girls do you think will buy that book there?' (Taraldsen 2002)

In well-formed subject extraction, the derivation will follow the steps in (23):

(23) a. [_λ T [_α subject ...]]
 b. [_γ *nP* [_λ T [_α subject ...]]]
 c. [_κ C [_γ *nP* [_λ T [_α subject ...]]]]
 d. [_σ subject [_κ C [_γ *nP* [_λ T [_α t ...]]]]]]

e. [subject [... [σ t [κ C [γ nP [λ T [α t ...]]]]]]]

In step (23b), an expletive is merged with λ in (20a) and (22a) while a resumptive pronoun is merged as Spec,TP in (21a). Taraldsen (2002) argues that *qui* is analyzed as *que + i*, which is an expletive-like element and is akin to the standard French *il* expletive. The merge applies externally and can be performed at the non-phase level. Thanks to the external merge, which yields overt Spec,TP, T can serve as a label and a labeling problem does not arise with λ or the T-headed set. Moreover, the merge does not induce the transfer of γ as it applies to the root at (23b) and is not counter-cyclic. As I have discussed, since C can label on its own, σ and κ can also be labeled even if the subject moves out from Spec,CP and turns into a copy. The proposed analysis can correctly predict that subject extraction from a full clause is not always ungrammatical. This not only supports the validity of the proposed analysis; it also argues that the generalization (1) is not correct.

Notice that the external merge cannot salvage subject extraction out of a full clause in A-movement. Consider (24):

(24) * The student seems [that [it is in the library]].

In the derivational process, the subject moves to Spec,CP to get out of the embedded clause as the TP or γ, which is the complement of the phase head, is cyclically transferred at the phase level. Notice that Spec,CP is an A'-position as C loses φ-features through feature-inheritance to T, and κ, with which the subject is merged, is not headed by a head bearing φ-features. From the embedded Spec,CP, it undergoes improper movement to the matrix Spec,TP, which is an A-position:

(25) [C [subject [... [σ t [κ C [γ nP [λ T_φ [α t ...]]]]]]]]
↑ ↑
A-movement A'-movement

Though transfer and labeling problems can be circumvented just as in (20)-(22), (24) will be ruled out for improper movement. The ill-formedness follows under the proposed analysis.

5. Consequences of the Proposed Analysis

The proposed analysis has three consequences. In this section, I discuss them in turn and show that they are theoretically and empirically favorable.

The first consequence is that literal deletion operations like Exfoliation and *that*-deletion can be wiped out and no extra mechanism beyond irreducible Merge is required to explain clause reduction. As I have discussed, external pair-Merge can reduce structure: T is adjoined to C and there is only one set above ν (i.e., a set headed by <C, T>) in the clausal spine. This consequence is desirable: deletion operations in syntax are extra mechanisms beyond irreducible Merge. Moreover, Exfoliation and *that*-deletion violate the No-Tampering Condition (NTC), one of the conditions that follow from general principles of minimal computation. NTC is stated as follows:

(26) Merge of X and Y leaves the two SOs unchanged. (Chomsky 2008:138)

Subject Extraction and Clause Size

Deletion operations remove a portion of the existing structure and change the structure created by Merge. Given NTC, the clause, once structured, cannot be restructured by deletion operations; when clause reduction occurs, the clause will be reduced from birth. External pair-merge of heads can remove a certain portion of the structure in accordance with (26).

The second consequence is that infinitival T can label. Recall that T is too weak to label alone, which is shown by (27a) and (28a):

- (27) a. * Which student do you believe [that is the most intelligent]?
 b. Which student do you believe [to be the most intelligent]? (=2a,c)
- (28) a. * The student seems [that is in the library].
 b. The student seems [to be in the library]. (=3)

However, (27b) and (28b) argue that labeling in the infinitival clause is not problematic even without overt Spec,TP, suggesting that infinitival T, unlike finite T, can label on its own, which has only been stipulated (see Epstein, Kitahara and Seely 2014). The labelability follows from my proposal on clause reduction. In (27b) and (28b), T is externally pair-merged to C and the amalgamated head <C, T> is created. <C, T>, which is externalized as *to* when it lacks tense, is syntactically on par with C. Since C is strong enough to serve as a label, <C, T> can label on its own without overt Spec. It follows from pair-merge of T to C that infinitival T is labelable.

The final consequence is that under-inheritance is explained by Merge. Legate (2011) (see also Mizuguchi 2014a, Ouali 2008) argues that when the subject *wh*-phrase moves to Spec,CP in such examples as (29), ϕ -features, as shown in (30), will not be inherited from C to T, with C retaining the features:

- (29) a. [Which student will read the book]?
 b. I wonder [which student will read the book].

(30) $[\sigma \text{ subject } [\kappa \text{ C}_\phi [\lambda \text{ T } [\alpha \dots t \dots]]]]$

In the proposed analysis of subject extraction, T is externally pair-merged to C and TP is reduced, which can bypass transfer and labeling problems, and the same argument applies to the derivation of subject movement within a single clause like (29): unless the clause is reduced, transfer and labeling problems will arise. Consider (31) and (32), and compare them with (8)/(9) and (15):

- (31) a. $[\text{---} [\kappa \text{ C } [\gamma \text{ which student } [\lambda \text{ T } [\alpha \text{ t } [\text{read-}v^* \text{ [the book]]]]]]]]]$ ($\gamma \rightarrow$ transferred)
 b. $[\sigma \text{ which student } [\kappa \text{ C } [\lambda \text{ T } [\alpha \text{ t } [\text{read-}v^* \text{ [the book]]]]]]]$ ($\lambda \rightarrow$ unlabeled)
- (32) $[\mu \text{ which student } [\delta \text{ <C, T> } [\alpha \text{ t } [\text{read-}v^* \text{ [the book]]]]]]]$ (\checkmark transfer; \checkmark labeling)

If the subject *wh*-phrase moves to Spec,TP, then the movement will transfer γ as it applies counter-cyclically, and the *wh*-phrase, which is embedded in γ , will be inaccessible and it cannot move to Spec,CP (=31a)). On the other hand, if the *wh*-phrase moves directly to Spec,CP, skipping Spec,TP, λ will not be labeled, with the result that the derivation violates Full Interpretation at the interfaces for the unlabeled SO (=31b)). These problems do not arise under (32).

Provided that pair-Merge applies and T is adjoined to C in the derivation of (29), T will not be available in the derivation; as a result, feature-inheritance does not take place and C is forced to keep ϕ . Under the proposed analysis, under-inheritance receives a Merge-based, hence principled explanation.⁷

Notice that under-inheritance by pair-Merge is supported by the halting of the ECM subject in the ECM infinitive. To see this, consider (33):

(33) The professor believes [the student to be the most intelligent].

In a series of his works, Lasnik (e.g., Lasnik 2004) persuasively argues that raising-to-object is optional and that the ECM subject can remain in the ECM infinitive. This suggests that the ECM subject moves only within the ECM complement. The movement creates XP-YP, which can only be labeled by agreeing heads. In (33), T is externally pair-merged to C and the amalgamated head $\langle C, T \rangle$ is created, which is externalized as *to*. Labeling of the XP-YP will be unproblematic if ϕ -features are under-inherited by the pair-merge. Thanks to under-inheritance, Label can locate ϕ -features in small *n* and $\langle C, T \rangle$ in the XP-YP created by the movement of the subject:

(34) The professor believes [_{μ} the student [_{δ} $\langle C_\phi, T \rangle$ [_{α} be [t [the most intelligent]]]]]

Since $\langle C, T \rangle$ bears ϕ and agrees with the ECM subject, the μ -marked set can be labeled $\langle \phi, \phi \rangle$ thanks to agreement or prominent feature sharing between the two heads.

To summarize, I have discussed three consequences of the proposed analysis of clause reduction. Notice that these consequences are brought by Merge, which is irreducible and forms the bare minimum part of UG, and endorse its role in FL.

6. Conclusion

In this paper, assuming SMT as the basic hypothesis in quest of an explanatory theory of language, I have considered the subject-extraction generalization proposed in the literature and addressed why clause size counts for subject extraction. I have claimed that Merge and Label, two elements of principled explanation, explain why subject extraction requires a smaller-than-full clause and that clause reduction for subject extraction follows from external pair-merge of heads, hence from

⁷ Mizuguchi (2014a) argues that feature-inheritance is preempted by head movement (i.e., internal pair-Merge). This proposal, however, leaves (a copy of) T in the derivation, making feature-inheritance to this T possible.

Under-inheritance of ϕ when T is pair-merged to C is also endorsed by Epstein, Kitahara and Seely (2012), who argue that ϕ -feature inheritance occurs for Case valuation by T.

Subject Extraction and Clause Size

Merge; the clause is reduced by nature. I have also proposed that C receives distinct externalizations depending on how it is merged. The discussion in this paper has shown that the relevant generalization receives a principled explanation under the assumption of SMT. Three consequences follow from the proposed analysis, which endorse the role of Merge in language.

The conclusion that has been reached through the discussion in this paper is that the simplest account is possible for subject extraction. The present paper is one illustration of language keeping strictly to SMT, with the basic hypothesis on language being supported.

References

- Bošković, Željko. 1997. *The Syntax of Nonfinite Complementation: An Economy Approach*. Cambridge, Mass.: MIT Press.
- Carstens, Vicki, and Michael Diecks. 2013. Parameterizing Case and Activity: Hyper-Raising in Bantu. In Kan, Seda, Claire Moore-Cantwell and Robert Staubs (eds.), *Proceedings of the Fortieth Annual Meeting of the North East Linguistic Society*. Amherst: University of Massachusetts, Graduate Linguistics Student Association, 99–118.
- Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.
- Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In Martin, Roger, David Michaels and Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass: MIT Press, 89–155.
- Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In Belletti, Adriana (ed.), *The Cartography of Syntactic Structures, Volume 3: Structures and Beyond*. Oxford: Oxford University Press, 104–131.
- Chomsky, Noam. 2007. Approaching UG from Below. In Sauerland, Uli and Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. Berlin: Mouton de Gruyter, 1–29.
- Chomsky, Noam. 2008. On Phases. In Freidin, Robert, Carlos P. Otero and Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. Cambridge, Mass: MIT Press, 133–166.
- Chomsky, Noam. 2010. Some Simple Evo Devo Theses: How True Might They Be for Language? In Larson, Richard K., Viviane Déprez and Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*. Cambridge: Cambridge University Press, 45–62.
- Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130: 33–49.
- Chomsky, Noam. 2015. Problems of Projection: Extensions. In Di Domenico, Elisa, Cornelia Hamann and Simona Matteini (eds.), *Structures, Strategies, and Beyond: Studies in Honour of Adriana Belletti*. Amsterdam: John Benjamins, 1–16.
- Chomsky, Noam. 2017. Language Architecture and Its Import for Evolution. *Neuroscience and Biobehavioral Reviews* 81: 295–300.
- Chomsky, Noam, and Howard Lasnik. 1993. The Theory of Principles and Parameters. In Jacobs, Joachim, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter, 506–569.
- Diercks, Michael. 2013. Indirect Agree in Lubukusu Complementizer Agreement. *Natural Language and Linguistic*

Manabu Mizuguchi

Theory 31: 357–407.

- Diesing, Molly. 1990. Word Order and the Subject Position in Yiddish. *Natural Language and Linguistic Theory* 8: 41–79.
- Engdahl, Elisabet. 1982. Restrictions on Unbounded Dependencies in Swedish. In Engdahl, Elisabet and Eva Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*. Stockholm: Almqvist and Wiksell International, 151–174.
- Erlewine, Michael Y. 2017a. Extraction and Licensing in Toba Batak. Ms. National University of Singapore, Singapore.
- Erlewine, Michael Y. 2017b. The Anti-Locality Signature of Quirks of Subject Extraction. Paper presented at workshop on quirks of subject extraction, National University of Singapore.
- Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely. 2012. Structure Building That Can't Be. In Uribe-Etxebarria, Myriam and Vidal Valmala (eds.), *Ways of Structure Building*. Oxford: Oxford University Press, 253–270.
- Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely. 2014. Labeling by Minimal Search: Implications for Successive-Cyclic A-Movement and the Conception of the Postulate “Phase.” *Linguistic Inquiry* 45: 463–481.
- Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely. 2016. Phase Cancellation by External Pair-Merge of Heads. *The Linguistic Review* 33: 87–102.
- Fukui, Naoki. 1993. A Note on Improper Movement. *The Linguistic Review* 10: 111–126.
- Halpert, Claire. 2015. *Argument Licensing and Agreement*. Oxford: Oxford University Press.
- Ishii, Toru. 2004. The Phase Impenetrability Condition, the Vacuous Movement Hypothesis, and *That-t* Effects. *Lingua* 114: 183–215.
- Lasnik, Howard. 2004. The Position of the Accusative Subject in the Accusative-Infinitive Construction. In Bhaskararao, Peri and Karumuri V. Subbarao (eds.), *Non-nominative Subjects, Volume 1*. Amsterdam: John Benjamins, 269–281.
- Legate, Julie. 2011. Under-Inheritance. Paper presented at the 42nd North East Linguistic Society, University of Toronto.
- Mizuguchi, Manabu. 2014a. Consequences of Feature Inheritance for Subject Movement. In Santana-LaBarge, Robert E. (ed.), *Proceedings of the Thirty-First West Coast Conference on Formal Linguistics*. Somerville, Mass: Cascadilla Proceedings Project, 325–334.
- Mizuguchi, Manabu. 2014b. Phases and Counter-Cyclicity of A-movement. In Park, Jong Un and Il-Jae Lee (eds.), *Proceedings of the 16th Seoul International Conference on Generative Grammar*. The Korean Generative Grammar Circle, 257–277.
- Obata, Miki, and Samuel D. Epstein. 2011. Feature-Splitting Internal Merge: Improper Movement, Intervention, and the A/A' Distinction. *Syntax* 14: 122–147.
- Ouali, Hamid. 2008. On C-to-T ϕ -Feature Transfer: The Nature of Agreement and Anti-agreement in Berber. In D'Alessandro, Roberta, Susann Fischer and Gunnar H. Hrafnbjargarson (eds.), *Agreement Restrictions*. Berlin: Mouton de Gruyter, 159–180.
- Pesetsky, David. 2016. Exfoliation: Towards a Derivational Theory of Clause Size. Paper presented at workshop on shrinking trees, Universität Leipzig.
- Taraldsen, Knut T. 2002. The *Que/Qui* Alternation and the Distribution of Expletives. In Svenonius, Peter (ed.), *Subjects, Expletives, and the EPP*. Oxford: Oxford University Press, 29–42.