SUPERIORITY EFFECTS IN MINIMALISM:
A CASE STUDY OF A-MOVEMENT

MANABU MIZUGUCHI

Dokkyo Medical University

This paper considers the Strong Minimalist Thesis proposed in Minimalist theory with a case study of superiority effects in A-movement. Given this thesis, Merge applies freely in syntax, which suggests that an NP can move over another NP. We argue that in derivations which result in superiority violations, Case features are transferred to the interfaces unvalued, claiming that interface conditions, not minimality, are responsible for the superiority effects. We also discuss two predictions of the proposed analysis, arguing that they are theoretically and empirically endorsed. Through our discussion, we show that syntax is only extrasyntactically constrained, concluding that language is perfectly designed.*

Keywords: superiority effects, minimality, free Merge, unvalued features, interface conditions

1. Introduction

Movement is a distinctive and unique property of human language and has been a primary focus of research up to the current Minimalist syntax. One of the most important discoveries that have been made in generative syntax is that movement is not free but constrained. This paper considers superiority effects in A-movement under the current Minimalism based on the Strong Minimalist Thesis (SMT), in which Merge is assumed to be syntactically unconstrained and applies freely. In the literature, a number of attempts have been proposed to explain superiority effects (say,
the Superiority Condition, Shortest Move, Attract Closest) but the crux of these proposals is to syntactically constrain the application of movement (i.e. Internal Merge): the movement of an NP over another NP violates minimality (locality) of movement/attraction.

The purpose of this paper is to reconsider this widely entertained, minimality-based explanation of superiority effects and propose a fresh analysis in SMT-based Minimalist theory. We propose that the movement of NPs is not intrasyntactically but extrasyntactically constrained by interface conditions imposed by the external systems, claiming that it is not minimality of movement/attraction but interface conditions that are responsible for superiority effects. Through our discussion in this paper, we show that principled explanation can be given to superiority effects in terms of the interfaces, concluding that the SMT is upheld.

This paper is organized as follows: in section 2, as our theoretical background, we first discuss a Minimalist view of language design and a structure-building operation Merge, in particular. Based on this background, in section 3, we argue that superiority in A-movement is not due to minimality of movement/attraction but to Case features transferred to the interfaces unvalued, proposing that the superiority is reducible to violations of interface conditions. In section 4, we consider two predictions of the proposed analysis and show that they are borne out. We conclude the paper in section 5 and briefly consider an implication for A′-movement superiority effects.

2. Theoretical Background

2.1. The Strong Minimalist Thesis as Minimalist Guidelines

One of the remarkable changes from Government and Binding (GB) syntax to Minimalist syntax is that the latter is based on the methodological desideratum of reducing the richness of Universal Grammar or UG (unexplained elements of S0) as much as possible and attributing the properties of language to UG-independent factors listed in (1) (Chomsky (2004: 106)):

(1) a. Interface conditions (the principled part of S0)
   b. General principles not specific to language (the third-factor principles)

The faculty of language (FL) (or the FL in the narrow sense), a key component of which is a computational system C_{HL} or syntax, is interfaced with the Conceptual-Intentional (CI) and Sensory-Motor (SM) systems. These language-external systems have their own conditions, imposing such conditions on the output objects generated by language so that they can interpret
and externalize them. General principles not specific to language (hence, not UG-specific principles as postulated in GB theory) are properties of nature and computational systems (like language) in particular, and include principles of computational efficiency, simplicity, elegance, etc. The current Minimalist theory assumes that to the extent that UG is simplified as much as possible and the properties of language keep to the independently motivated factors given in (1), principled explanation deeper than explanatory adequacy will be achieved for human language, and the FL will be a “perfect” system, meeting interface conditions in a way satisfying the third-factor principles of minimal/simple computation. This hypothesis, called the “Strong Minimalist Thesis” (SMT), is the backbone of Minimalist theory and is also a hallmark which prominently distinguishes it from the other syntactic frameworks.

2.2. Free Merge

With this language design in mind, we now focus on structure building in SMT-based Minimalism, which is crucially relevant to movement. Provided that any generative/recursive system must be equipped with a compositional procedure to construct objects, the FL should have an operation that creates structures (or expressions) with lexical items. Minimalist theory assumes the operation Merge as an irreducible and bare minimum part of UG (a “virtual conceptual necessity”), eliminating from the grammar phrase structure rules and X-bar theory. Merge, a simple set-formation operation, recursively takes $n$ syntactic objects (SOs) already formed and creates a new SO (i.e. an $n$-membered set) out of them. Without any further assumptions, Merge is unconstrained and applies freely. However, given the third-factor principles, which yield minimal/simple computation, $n$, most reasonably, reduces to two, with the result that Merge is binary, and two SOs merged are left unchanged (the No-Tampering Condition), with no new objects (indices, traces, bar levels, labels) added in the process of Merge (the Inclusiveness Condition). Thus, Merge, most simply, is executed as in (2):

\[
\text{(2)} \quad \text{Merge}(\alpha, \beta) = \{\alpha, \beta\}
\]

Phrase structure created by the simplest execution of Merge illustrated in (2) (call it simplest Merge) is bare in that it is free from linear order, labels or endocentricity (projection), all of which are stipulated properties of phrase structure and which can be assigned independently when SOs are mapped to SEM and PHON through Transfer (Chomsky (2007, 2008, 2013)).

Under the assumption of the SMT, anything can be merged with anything and Merge applies freely to any two SOs. Selectional relations are prop-
erties of the CI system, with the well-formedness of the relation between α and β subject to CI interface conditions (as argued in, say, Chomsky (2004), Fortuny (2008); but see Adger (2003), Pesetsky and Torrego (2007) among others for different proposals). As a further consequence of unconstrained Merge, we argue that a feature that drives Merge such as the edge-feature (EF) (Chomsky (2007, 2008)) as well as the assumption that only a lexical item can bear EF to be merged with an SO (Narita (2011)) turns out to be a mere stipulation that should be dispensed with. Given the methodological desideratum of Minimalist theory, no constraint will be imposed on the application of Merge. Hence, we submit that Merge, under the SMT, is Merge-α (Merge anything anywhere anytime): it applies freely without caring about syntactic objects to be derived, as far as it abides by the third-factor principles of minimal/simple computation; the convergence/well-formedness of derived objects is determined by the CI and SM systems through interface conditions they impose after derivations are sent to these systems through Transfer. The application of Merge is only extrasyntactically constrained by (1), with only certain merge choices converging at the interfaces. In this simplest conception of phrase-structure building, syntax is not crash-proof and generates, in principle, an infinite number of syntactic objects that are unusable (i.e. cause crash at the interfaces) or are assigned interpretations with various degrees of deviance (Chomsky (2008)) (see also Ott (2010) for much related discussion).1

2.3. Movement as Internal Merge

One of the important consequences of free, unconstrained Merge or Merge-α is that the operation will apply internally inside a phrase as well as externally, with the result that movement can be reformulated as one form of Merge (i.e. Internal Merge, IM) (Chomsky (2004) et seq.). Moreover, if movement is nothing more than Merge and Merge is intrasyntactically unconstrained, IM is as free as External Merge (EM). Then, unlike what has

1 A reviewer wonders if free Merge will increase the overall computational burden as the number of derivations increases that would crash. We argue that as far as syntax operates in the simplest manner, abiding by the third-factor principles, it can generate derivations or SOs ad infinitum without the increase of computational burden, whether they converge or crash; it is the job of the interfaces to decide the convergence or crash of derived SOs, which does not matter for syntax. As another reviewer points out, however, we cannot calculate exactly the complexity of the computational system unless explicit standards are set for the calculation. The topic is beyond the scope of this paper and we leave further clarification of this problem for future research.
been commonly assumed, movement does not require any stipulated trigger at all (say, the EPP/OCC feature, EF); nor is it subject to the Last Resort (cf. Chomsky (1986)).

In this section, we have reviewed a Minimalist hypothesis on language design and a structure-building operation Merge, in particular, and argued that (Internal) Merge is intrasyntactically unconstrained but is extrasyntactically constrained by interface conditions imposed by the interface systems and the third-factor principles. In the next section, we will explore SMT-based Minimalism with superiority effects in A-movement.

3. Superiority in A-Movement

3.1. Shortest Move/Attract Closest

It has been observed that A-movement shows superiority effects. Consider (3) and (4):

(3) a. Ann was given a record (by Debbie).
   b. *A record was given Ann (by Debbie).

(4) a. Robert was sent a telegram (by Wayne).
   b. *A telegram was sent Robert (by Wayne).

(Stowell (1981: 325))

In these examples, an indirect object (Obj indirect) can be A-moved (raised) while a direct object (Obj direct) cannot. Suppose that the double object construction has the structure illustrated in (5), where double objects require two vP phases (each of these phase heads bearing unvalued phi(ϕ)-features) for Case valuation, with the phi-features being inherited by Vs (Chomsky (2007, 2008), Epstein et al. (2012), Richards (2007)); Obj indirect is merged with lower vP, and hence merged higher than Obj direct (in (5) and elsewhere, right brackets are omitted for simplicity):

(5) \[ v_P \text{Subj} [v_P [v_P V_{\phi}] [v_P \text{Obj indirect} [v_P [v_P V_{\phi}] \text{Obj direct}]...]] \]

(order irrelevant)

With (5) in mind, it can be considered that the superiority effects are ex-

\[ \text{search}\]

2 It has often been argued that IM, unlike EM, requires the operation Search, by which a head looks for an element which undergoes IM to its edge. Under the free-Merge hypothesis, IM is free from search. Search by a head implies selection in that the head looks for a particular target, constraining IM. Thus, postulating search as a property of IM is unfavorable to free Merge. IM, as we have argued, is not triggered by any feature under free Merge and unlike Agree, probe-goal is not established and no search is required. That search is irrelevant to IM is a desirable consequence if EM and IM are both instantiations of the same operation (= (2)).
explained by minimality of movement/attraction: of the two NPs, an NP which moves is closer to the TP edge (Shortest Move) or to T (Attract Closest) than the other. Minimality (of movement/attraction) is attributable to the third-factor principles of minimal/simple computation and the appeal to minimality to constrain movement may seem a reasonable approach. This minimality-based explanation, however, is unavailable in the SMT-based view of structure building. Under the free-Merge hypothesis spelled out in the last section, no constraint will be imposed on the application of Merge in syntax (except that Merge applies most simply, as shown in (2)) and if movement is IM, any NP can move to the TP edge thanks to free Merge and locality is irrelevant; a lower NP can be internally merged over a higher NP without any problem. Moreover, as we have discussed in section 2 as well as in footnote 2, no triggering feature is postulated under free Merge and IM is free from search: probe-goal cannot be resorted to in order to constrain the IM of NPs through minimal search. Given the free-Merge hypothesis, the application of IM is independent of probe-goal relations, even if NPs enter into probe-goal relations with T (which is empirically endorsed by the there-expletive construction) (cf. Chomsky (2000)).

3.2. Superiority as Violations of Interface Conditions

We argue that contra what has been proposed, what causes superiority effects in (3) and (4) is not the IM of a lower NP over a higher NP, which is allowed by syntactically unconstrained Merge, and hence that minimality of movement/attraction is irrelevant to the effects. We claim that the ill-formedness of (3b) and (4b) receives a principled account in terms of the interfaces (= (1a)). Take (3b) as our example. In (3b), the indirect object Ann agrees with higher \( V_{\{\phi\}} \) and has its Case feature valued in a higher \( v^*P \) phase while the direct object a record agrees with T\( _{\{\phi\}} \) and its Case feature is valued in a CP phase. Thus, lower \( v \) does not bear phi-features and is a weak phase. The structure is illustrated as follows:

\[
(6) \quad [v^*P \ v^* [VP \ V_{\{\phi\}} [v \ [VP \ V \ OBJ_{direct}]]]]
\]

Assuming that Transfer applies cyclically at the strong phase level, in (6), the complement of lower \( v \) is not cyclically transferred but that of higher \( v^* \) is. This requires the IM of OBJ\textsubscript{direct} to the edge of higher \( v^*P \); otherwise, it would be trapped in lower VP, which is transferred as part of higher VP at the higher \( v^*P \) phase level (in (7) and elsewhere, SOs in angled brackets \(<SO>\) are intended as copies (i.e. lower occurrences of the same lexical item created by IM)): 
The derivation continues and a CP phase is created by the merge of T and C on top of higher \( \nu^*P \), with the phi-features of C being inherited by T:

\[
(8) \quad [\text{CP C} \ [TP T_{\phi}] \ [\nu^*P \ \text{Obj}_{\text{direct}}] \ [\nu^* \ [\text{VP V}_{\phi}] \ [\nu P \ \text{Obj}_{\text{indirect}}] \ [\nu \ [\text{VP V} <\text{Obj}_{\text{direct}}>].]
\]

In this CP phase, \( T_{\phi} \) agrees with \( \text{Obj}_{\text{direct}} \) and the phi-features as well as the Case feature of the object are successfully valued. \( \text{Obj}_{\text{direct}} \) is internally merged to the edge of TP (Epstein et al. (2012), Narita (2011); see also Chomsky (2008)) and the following SO is finally derived:

\[
(9) \quad [\text{CP C} \ [TP \ \text{Obj}_{\text{direct}} \ [T_{\phi}] \ [\nu^*P <\text{Obj}_{\text{direct}}>] \ [\nu^* \ [\text{VP V}_{\phi}] \ [\nu P \ \text{Obj}_{\text{indirect}}] \ [\nu \ [\text{VP V} <\text{Obj}_{\text{direct}}>].]]
\]

This derivation (which is also an instance of so-called “improper movement”) incurs ill-formedness as shown in (3b). Since Merge applies most simply, abiding by the third-factor principles, unconstrained Merge can produce the derivation illustrated in (9). Given the SMT, the SO in (9) will be ruled out by the interfaces. We argue that the ill-formedness caused by the derivation is reducible to the fact that features are sent to the interfaces unvalued. To see this, consider once again the derivational process in which \( \text{Obj}_{\text{direct}} \) is internally merged from its externally merged position to the edge of higher \( \nu^*P \) for cyclic transfer of higher VP (= (6), (7)). \( \text{Obj}_{\text{direct}} \) bears an uninterpretable Case feature, which is to be valued via its phi-feature agreement with \( T_{\phi} \) in a CP phase. This suggests that when the object is internally merged, the IM creates a copy with an unvalued Case feature and the two occurrences of the object (not only the copy in the higher \( \nu^*P \) edge but also the one in the externally merged position) bear unvalued Case features. This derivation will cause crash at the interfaces because the copy in the externally merged position is sent to the interfaces with its Case feature unvalued when higher VP is transferred at the higher \( \nu^*P \) phase level (“u” before features means “unvalued”):

\[
(10) \quad \begin{align*}
a. \quad & [\nu^*P \ \nu^* \ [\text{VP V}_{\phi}] \ [\nu P \ \text{Obj}_{\text{indirect}}] \ [\nu \ [\text{VP V} \ \text{Obj}_{\text{direct}}{u.Case}].] \\
& \quad \text{Transfer} \quad \Rightarrow \quad \text{CRASH}
\end{align*}
\]

We should note that the Case-feature valuation of the occurrence in the higher \( \nu^*P \) edge through its phi-feature agreement with \( T_{\phi} \) in a CP phase cannot render the Case feature of the lower copy valued: by the time the agreement takes place, the copy in question has already been transferred
with its Case feature unvalued.³

As we have argued, the derived structure (7) (= (10b)) violates interface conditions banning unvalued (that is, unspecified) features when it is transferred, and is ruled out externally by the interfaces. Superiority effects in A-movement are reducible to the factor (1a). (3b) is ill-formed, even though syntax, conforming to the third-factor principles, performs minimal/simple computation to derive (9) as its SO. The same explanation applies to (4b).

Let us now turn to the well-formed examples (3a) and (4a). To the extent that our proposal is on the right track, no violations of interface conditions are incurred here. We take (3a) as our example. In this case, the indirect object Ann agrees with $T_{\{\phi\}}$ and its Case feature is valued in a CP phase; the direct object a record, on the other hand, agrees with lower $V_{\{\phi\}}$ and its Case feature is valued in a lower $v^*P$ phase, with higher $v$ bearing no phi-features and being a weak phase:⁴

\[
(11) \quad [vP \ v \ [vP V [vP \ Obj_{\text{indirect}} [V^* [VP V \ Obj_{\text{direct}}]].]]
\]

The derivation continues and the merge of $T$ and $C$ on top of (11) generates a CP phase:

\[
(12) \quad [CP \ C \ [TP T_{\{\phi\}} \ [vP V [V^*P \ Obj_{\text{indirect}} [V^* [VP V_{\{\phi\}} \ Obj_{\text{direct}}]].]]
\]

In (12), the IM of $Obj_{\text{indirect}}$ to the edge of TP does not cause any problem at the interfaces: unlike in (3b), where $Obj_{\text{direct}}$ is internally merged to the edge of higher $v^*P$ and its copy in its externally merged position bears an unvalued Case feature, $Obj_{\text{indirect}}$ agrees with $T_{\{\phi\}}$ and has its Case feature valued before it is internally merged to the edge of TP. This is to say that

³ A reviewer has noted the possibility that the Case feature of a direct object is valued by multiple Agree with higher $v^*$ in (10a). We argue that this possibility is excluded in agreement languages like English: in such languages, where there is a one-to-one agreement relation between probes and goals as discussed in, say, Fukui (1986) and Kuroda (1988), phi-feature agreement with an indirect object values the phi-features of higher $v^*$, which inactivates the probe and blocks further agreement.

⁴ We argue that in the double object construction, a passive morpheme can be attached to either one of the two $v^*$s, absorbing phi-features (and a theta property (cf. Jaeggli (1986))) of the $v^*$ it is attached to and making phi-feature agreement for Case valuation impossible. Since only one passive morpheme is allowed in passives, which is evidenced by the impossibility of multiple passive morphemes like *given en* (*give* + -en + -en), the absorption only takes place either at higher $v^*$ as in (11) or at lower $v^*$ as in (6), but not at both. I thank a reviewer for directing me to this discussion.
the Case feature of its copy (the lower occurrence of Obj\textsubscript{indirect}) is transferred to the interfaces valued. Consequently, movement from the lower \(v^*P\) edge to the TP edge does not violate interface conditions banning unvalued features at the interfaces. The structure ultimately derived (= (13)) is ruled in as well-formed at the interfaces and the derivation converges:

\[
\begin{align*}
(13) \quad & [CP \ C: [TP \ Obj_{\text{indirect}\{\text{Case}\}}] \ [T_{\{\phi\}}] \ [v^*P \ v \ [vP \ [v^*P <\text{Obj}_{\text{indirect}\{\text{Case}\}}> \ [v^*P \ \text{Transfer} \ \rightarrow \ \text{CONVERGE} \ [vP \ [v_{\{\phi\}} \ \text{Obj}\text{direct}]..]]] \]
\end{align*}
\]

The well-formedness of (4a) is explained likewise.\(^5\)

Another example of superiority effects in \textit{A}-movement is illustrated by so-called superraising. Consider (14):

\(^5\) A reviewer has pointed out that the grammatical asymmetry in (3) and (4) is not universal, wondering how this cross-linguistic variation could be explained. We argue that in languages like Greek, Icelandic and Japanese, where a direct object can be passivized, it can be merged higher than an indirect object: that is, (i) as well as (5) can be generated by EM:

\[
(\text{i}) \quad [vP \ \text{Subj} \ [v^*P \ [vP \ [v^*P \ \text{Obj}\text{direct} \ [v^*P \ [vP \ \text{Obj}\text{indirect}]..]]] \]
\]

See Doggett (2004) and references cited therein for evidence supporting (i), and the correlation between the availability of direct object passivization and the existence of (i). Given (i), Obj\text{direct} can agree with T\textsubscript{\{\phi\}} before its movement to the TP edge and no copy with an unvalued Case feature is left behind by the IM. We assume that (i) is possible for languages when, as argued and demonstrated in Doggett (2004), Obj\text{indirect} is syntactically expressed as PP.

A more recalcitrant case is found in British English, where only (5) can be produced in active sentences (= (ii)) but still, a direct object can be passivized (= (iii)):

\[
\begin{align*}
(\text{ii}) \quad *\text{Mary gave the book the professor.} \\
(\text{iii}) \quad \text{A book was given John and Mary.}
\end{align*}
\]

We follow Doggett (2004) and submit that Obj\text{direct} \textless Obj\text{indirect} merger by EM is possible in British English only when Obj\text{direct} is passivized. There are in fact languages where passivization (or raising) renders otherwise ill-formed active sentences well-formed. For instance, in languages like Chichewa, certain instrumental objects are allowed when raised to the subject position; however, they cannot appear in the object or first-merged position. Consider (iv):

\[
\begin{align*}
(\text{iv}) \quad & \text{a. Khásu li-ma-(li)-lim-its-idw-a chi-manga ndí Jóni.} \\
& \text{höe it-Habit-(it-)farm-Cause-Pass-Indic corn by John} \\
& \text{‘The hoe is used by John to farm corn with.’} \\
& \text{b. *Jóni á-ma-(yi-)lemb-éts-a péní.} \\
& \text{John he-Habit-(it-)write-Cause-Indic pen} \\
& \text{‘John writes with a pen.’}
\end{align*}
\]

(Marantz (1984: 143), cited in Doggett (2004: 98))

We argue that direct object passivization like (iii) is one illustration of this still unclear syntactic phenomenon, passivization allowing otherwise ill-formed Obj\text{direct} \textless Obj\text{indirect} merger. Further exploration has to be left for future.
(14) *John seems that it was told <John> that his mother was beautiful.

In the derivation, the subject John, which is externally merged as an internal argument of tell, moves over the expletive to the edge of the matrix TP. As in (3b) and (4b), this movement is possible under the free-Merge hypothesis adopted in the current Minimalist syntax; however, the derivation is ruled out by the interfaces because the IM of the subject to the edge of the embedded CP for cyclic Transfer yields a copy with an unvalued Case feature in its externally merged position, which is transferred to the interfaces unvalued when the embedded TP is transferred at the embedded CP phase level: the Case feature of John is valued through the agreement of the subject and T_{(4)} in the root CP phase and it is still unvalued when the subject is internally merged to the embedded CP edge:

(15) [CP John_{u,Case} [C [TP it was told <John_{u,Case}> that his mother was beautiful]]]

As we have demonstrated, superraising is explained on par with (3b) and (4b) by the creation of a crash-causing copy (i.e. a copy with an unvalued Case feature) through IM.

In this section, we have argued that the ill-formed examples of A-movement are ruled out by the fact that the IM of an NP creates a copy with an unvalued Case feature in its externally merged position, which is to be transferred unvalued. As we have demonstrated, syntax, conforming to principles of minimal/simple computation, can generate SOs in which a lower NP can move over a higher NP; however, the derived SOs, when transferred, contain unvalued features and are ruled out extrasyntactically by the interfaces because such features violate interface conditions banning unspecified features at the interfaces and crash the derivation.

The proposed analysis of superiority effects in A-movement is also superior to an account of (3b), (4b) as well as (14) in terms of improper movement (whatever it is in Minimalism). Before we leave this section, we discuss an improper-movement-based account. As we have noted, the movement involved in the derivations of (9) and (15) is also an instance of improper movement: an NP is moved from an A-position (a first-merged, theta-position) to an A′-position (a phase edge), from which it is moved back to an A-position (the TP edge). In the discussion of the ill-formed examples, we should consider another possible derivation of the examples under SMT-based syntax, which will be permitted under traditional improper
movement but correctly excluded under our explanation of superiority effects. Suppose that the object (the derived subject) is a *wh*-phrase and is internally merged to the CP edge from the edge of higher v*P/embedded CP without being internally merged to the TP edge (contra Chomsky (2008)). The movement of a subject *wh*-phrase from the v*P/CP edge in this way is theoretically unproblematic: phi-features inherited by T can be valued by Agree without establishing a Spec-head relation; furthermore, as we have discussed in section 2, T does not have any feature triggering movement (say, the EPP/OCC feature, EF) under free Merge. Also, there are empirical arguments against the movement of a subject *wh*-phrase to the TP edge in subject *wh*-movement, with the edge position not being created (e.g. Mizuguchi (2013) and references cited therein). The movement, which is from the v*P/CP edge (A’-position) to the CP edge (A’-position), is considered a case of proper movement. Then, the prediction is that the *wh*-counterparts to (3b), (4b) and (14) will be well-formed under traditional improper movement. This prediction, however, is not borne out. Consider (16):

(16) a. *What was given Ann (by Debbie)?
    b. *What was sent Robert (by Wayne)?
    c. *Who seems that it was told that his mother was beautiful?

The ill-formedness, on the other hand, can be correctly explained by the proposed analysis of superiority in A-movement in terms of unvalued features at the interfaces: the internal merge of an object to the higher v*P/embedded CP edge creates a copy with an unvalued Case feature in its first-merged position, which is to be transferred unvalued, whether the object is a *wh*-phrase or not. The ill-formedness of (16) is explained in the same way as that of (3b), (4b) and (14), and the examples are also ruled out correctly by the interfaces under SMT-based syntax.

The analysis we have proposed can not only explain improper movement; it is also superior in that it can explain examples which cannot be accounted for by assuming improper movement.

4. Predictions

If superiority in A-movement is due to the fact that IM creates a copy with an unvalued Case feature, which is transferred to the interfaces unvalued, then we get two predictions. In this section, we discuss them in turn.
4.1. Superraising in Moroccan Arabic

The first prediction is that if an NP has its Case feature valued before its A-movement, superiority effects will not be observed even if it is internally merged over another NP. This prediction is indeed borne out by superraising in languages such as Moroccan Arabic. Consider the following examples (see Ura (1994) for more examples of superraising):

(17) a. Ttshab-et-li mmi [belli šaf-Ø-ha seemed-3Sg-F-to-1Sg mother-1Sg Comp saw-3SgM-3SgF muhend <mmi> fsefru ]. Mohand in-Sefrou

b. Ttshab-li [belli šaf-Ø-ha muhend seemed-3Sg-to-1Sg Comp saw-3SgM-3SgF Mohand mmi fsefru ]. mother-1Sg in-Sefrou

‘It seemed to me that Mohand saw my mother in Sefrou.’ (Ura (1994: 10))

In (17a), the embedded object mmi agrees with the matrix verb ttshab and raises to the surface subject position (i.e. the TP edge) in the matrix clause. That the NP is in the TP edge is also suggested by the order between the verb and the NP. Ura (1994) reports that in matrix clauses, the subject usually follows the predicate in Moroccan Arabic. In (17b), on the other hand, the expletive pro is merged in the matrix subject position, with the verb agreeing in third person singular masculine. From these observations, it can be assumed that mmi has A-moved over the other NP muhend to the matrix subject position from the embedded clause to derive (17a).

The well-formedness of (17a) is correctly predicted by the proposed analysis. In the derivation of (17a), mmi has to move through phase edges in the course of the derivation for cyclic Transfer. Unlike in (3b), (4b) and (14), however, since the NP has its Case feature valued in its first-merged position thanks to its phi-feature agreement with V in the embedded \( v^*P \), which is morphologically evidenced by šaf-Ø-ha, its IM does not yield copies with unvalued Case features and the derivation does not crash even if TP and VP are cyclically transferred:

(18) \[ CP mmi_{\text{Case}} [belli \begin{array}{c} TP \ šaf-Ø-ha \ muhend \ [v_\text{P} \ <mmi_{\text{Case}}>] \ \\
\text{Transfer} \rightarrow \text{CONVERGE}
\end{array}] \]

\[ <mmi_{\text{Case}}> \ fsefru] \]

In other words, the well-formedness of (17a) is explained on par with the
well-formedness of (19); as shown in (20), the Case features of the subject *John* and the object *what* are valued via Agree in their first-merged positions before they are internally merged:

(19) What did John buy?

(20) a. \([v^*P \text{what}_{\text{Case}} \text{[John } [v^* [v_P \text{buy}_{\phi}] \text{<what}_{\text{Case}}>..]]\]

b. \([CP \text{what} [C [TP \text{John}_{\text{Case}}] [T_{\phi} [v^*P \text{<what}> \text{[<John}_{\text{Case}}]> [v^* [v_P \cdots ..]]\]

Superraising in Moroccan Arabic argues for the proposed analysis of superiority effects in A-movement; it also shows that because trans-phasal movement must proceed via phase edges, “improper” (that is, A-A’-A) movement can be executed in language and that A-A’-A movement steps are not the cause of ill-formedness in ill-formed cases of improper movement.6

4.2. Long-Distance A-Movement

Our interface-based explanation of superiority in A-movement predicts that in long-distance A-movement such as (21), the subject does not move through the edges of raising TPs:

(21) John seems to be likely to be in the library.

---

6 The question arises why the English counterpart to (17a) is ill-formed:

(i) *The bagel seems (that) John ate.

The ill-formedness of *The bagel* is valued in situ before it is internally merged to phase edges.

The ill-formedness of (i) can be explained if we assume, following Carstens (2011), that gender works as an infinitely reusable Activity feature in some languages. We argue that in superraising languages, a gender feature is syntactically available and that NPs can recursively agree with the matrix T in phi-features for the reusable Activity feature, which accounts for (17a) and superraising in other languages discussed in Ura (1994). The availability of a gender feature for further agreement is also suggested by the fact that the multiple subject construction is possible in superraising languages (Ura (1994)): a non-canonical subject in the construction can engage in further agreement with T thanks to a reusable gender feature. In English, on the other hand, we argue that a gender feature cannot be used to make NPs available for recursive phi-feature agreement; only a Case feature, which is a finitely usable Activity feature, is available to make NPs active for agreement. Consequently in (i), the matrix T cannot agree with *the bagel* in phi-features and the features will be left unvalued in the derivation; they will violate interface conditions banning unvalued features when transferred, causing crash at the interfaces.

The suggested analysis falls under the Uniformity hypothesis assumed in Minimalism: syntax is a computational system with a minimal set of invariant operations, with parameterization reduced to independently motivated, restricted morpho-featural cross-linguistic variation in the lexicon (Chomsky (2008), Epstein et al. (2013), Fukui (1986) and references cited therein). The ill-formedness of (i) in contrast with the well-formedness of (17a) can follow independently in our proposal from parameterization in the lexicon.
In the derivation of (21), the subject *John* agrees with $T_{\{\phi\}}$ in the root CP phase. The successive-cyclic movement of the subject would create copies with unvalued Case features in the edges of embedded TPs. The copies would be transferred with their Case features unvalued when the matrix TP is shipped off to the interfaces, causing crash. To the extent that our proposal is on track, it follows that long-distance A-movement is a one-step movement from the first-merged position to the edge of TP when the Case feature of the subject has been valued in situ via long-distance agreement.

In the literature, there are arguments for and against successive-cyclic A-movement (Bošković (2002), Castillo et al. (2009), Esptein and Seely (2006), Grohmann et al. (2000), Lasnik (2003) among others). In terms of the current Minimalist theory, the prediction is theoretically endorsed by the fact that successive-cyclic A-movement produces an SO which causes an interpretive problem at the CI interface. In phase-based syntax, IM occurs in simultaneity with Transfer at the phase level so that the occurrences created by IM can be identified as copies (i.e. multiple occurrences of the same lexical item) by the interfaces under the assumption of the Inclusiveness (Chomsky (2007) et seq.); otherwise, they would be identified as repetitions (distinct occurrences of the same lexical item) because the interfaces cannot determine whether the occurrences created are by IM or by EM. In (21), since Transfer applies only at the root CP phase level, if A-movement is successive cyclic and IM is executed before Transfer, the occurrences of the subject created in intermediate positions would be identified as distinct SOs (repetitions) at the interfaces when transferred, yielding distinct/non-chain interpretations, which would incur a violation of Full Interpretation (hence, cause deviance) and the derivation would be extrasyntactically ruled out by the CI interface.

On the empirical side, arguments for successive-cyclic A-movement, as persuasively discussed in Castillo et al. (2009), Epstein and Seely (2006) and Grohmann et al. (2000), among others, can be subject to explanations not relying on intermediate copies in the edges of raising TPs. For instance, consider reflexive binding in (22) and quantifier float in (23):

(22) *John* seems to Mary to appear to *himself* to be happy.

(23) The boys seem (all) to appear (all) to (all) like ice cream.

As for (22), suppose that transferred TP defines a binding domain for reflexives. Then, *John* can bind *himself* in the matrix TP without positing a copy in the edge of *to appear* clause. Moreover, if *himself* in (22) is actually a logophor, as claimed by Castillo et al. (2009: 95), which is suggested by the non-complementary distribution of pronouns and reflexives in
(24), then the fact that himself refers to John in (22) is attributable to this property, and not to a copy created by successive-cyclic A-movement (see references cited above for more discussion of data and alternative solutions):

(24) It seems to John to appear to him/himself that the earth is flat.

Likewise, an alternative account is available for quantifier float in (23). It has been assumed that floating quantifiers indicate positions of intermediate copies of movement. Since the quantifier all can float in (23), it may argue for successive-cyclic A-movement. Quantifier float a la Sportiche (1988), however, is not the only way to explain the phenomenon. Bobaljik (1995) argues that floating quantifiers are adverbs and modify the predicate in a predictable manner with respect to some NP. If quantifiers are adverbs, then they can float irrespective of successive-cyclic A-movement and can surface in positions in which copies of A-movement would not be expected but adverbs do appear. This is in fact borne out by (25):

(25) a. The magicians (all) should (all) have (all) arrived before the show begins.
   b. The votes (all) will (all) have (all) been counted by midnight.

These examples, cited from Bobaljik (1995: 206), show that the quantifier all appears in positions not created by successive-cyclic A-movement. This persuasively argues for floating quantifiers as adverbs and speaks against a Sportiche-style analysis of quantifier float.

Finally, Chomsky (2008) provides (26) as an argument for successive-cyclic A-movement:

(26) Of which car is the driver/picture likely to cause a scandal?
(27) *Of which car did the driver/picture cause a scandal?

(Chomsky (2008: 153))

Suppose that the subject-island effect is caused by the fact that a wh-phrase is moved out of an NP in a phase edge. (27) is ill-formed because the wh-phrase is extracted out of the subject the driver/picture in the v*P edge at the CP phase level. Chomsky argues that if the subject moves successively cyclically through the edges of raising TPs, of which car can be extracted from the non-phase edge position in the course of the derivation (= (28)). Since the extraction, unlike in (27), is not from the phase edge (the v*P edge), the subject-island effect is not incurred:

(28) [CP [C [TP is likely [TP [NP the driver/picture of which car], [to [v*P t₁ cause a scandal].]]]]

We argue that the well-formedness of (26) can also be subject to an ex-
planation without assuming local steps. Mizuguchi (2009), who proposes a phase-based analysis of licit adnominal PP extraction like (29), argues that *wh*-movement is unproblematic in (29) because *of which article* is not externally merged in the subject but with VP (i.e., it is adjoined), from which it is internally merged to the CP edge (= (30)), with the *wh*-phrase being interpreted as part of the subject not through movement but through mutual c-command via cyclic Transfer: the VP-adjoined *wh*-phrase c-commands the subject *a review* in VP, which in turn c-commands the *wh*-phrase when moved to the TP edge:

(29) Of which article did a review come out yesterday?

(30) \[
\begin{array}{c}
\text{CP} \\
\text{[C \ [TP \ [NP \ a \ review] \i \ [T \ [VP \ [VP \ come \ out \ t_i \ yesterday] \text{of \ which \ article]..]]}
\end{array}
\]

Thus, the *wh*-phrase is not extracted out of the subject. We refer the reader to Mizuguchi (2009) for details (see also Broekhuis (2013) and van Craenenbroeck and den Dikken (2006) for the argument that adnominal PPs are moved from NP/subject-external positions).

Given this adjunction analysis of adnominal PPs, the well-formedness of (26) follows without assuming successive-cyclic A-movement. In (26), the raising predicate provides VP to which *of which car* can be adjoined. This enables the *wh*-phrase to c-command the subject *the driver/picture* in the *v*P edge, which in turn c-commands the VP-adjoined *wh*-phrase when raised to the matrix TP edge. Since the two c-command each other when TP is transferred at the CP phase level, the interpretation is well-formed at the CI interface. Importantly, the analysis does not assume extraction out of the subject in a non-phase edge position. (27), on the other hand, is ill-formed because *of which car*, which is adjoined to the *cause*-headed VP, cannot c-command the subject in the *v*P and TP edges, and mutual c-command between *of which car* and the subject cannot be established at any point of cyclic Transfer.

The proposed explanation is endorsed by the fact that (26) will turn ill-formed if the preposition is stranded and the *wh*-phrase alone is moved. Consider (31):

(31) *Which car is the driver/picture of likely to cause a scandal?*

The ill-formedness of (31) is correctly predicted because *which car*, unlike the prepositional phrase *of which car*, cannot be adjoined and has to be moved out of the subject, which causes the subject-island effect if the subject does not move successive cyclically and the extraction is forced from
the v*P edge. Chomsky’s analysis, on the other hand, would predict that (31), along with (26), is well-formed because the subject moves successive cyclically through TP edges and the wh-phrase can be extracted from the subject in the non-phase edge of raising TP.

The examination of the three kinds of evidence for successive-cyclic A-movement suffices to show that the alleged empirical evidence can not only receive alternative explanations without intermediate copies, but that it is also weak in firmly endorsing successive-cyclic A-movement when a wider range of data is taken into account. From the theoretical and empirical discussions in this section, we can conclude that long-distance A-movement is one-fell swoop, which supports the prediction of the proposed analysis.

5. Conclusion

In this paper, we have considered superiority effects in Minimalist theory with A-movement. Given SMT-based Minimalism, especially the free-Merge hypothesis, we have claimed that superiority violations in A-movement are reducible to the fact that Case features are transferred unvalued, causing crash at the interfaces. Thus, interface conditions, not minimality of movement/attraction, are responsible for the superiority effects. We have also discussed two predictions of the proposed analysis, showing that they are theoretically and empirically upheld. To the extent that our discussion is correct, principled explanation can be given to superiority effects in terms of the interfaces and the SMT is endorsed.

The analysis proposed in this paper has an implication for A’-movement.7 It has been noted that A’-movement, say wh-movement, also shows superiority effects:

(32)  a. Who bought what?
      b. *What did who buy?

To the extent that our proposal in this paper is correct, (32b) can be generated by free Merge without any problem and minimality of movement/attraction is also irrelevant to superiority effects in A’-movement. In fact, Mizuguchi (2014) persuasively argues under the free-Merge hypothesis that A’-movement superiority effects like (32b) and others are attributable to violations of interface conditions imposed by the CI and SM systems, claiming

7 I thank a reviewer for directing me to the discussion of superiority effects in A’-movement.
that the effects are reducible to the UG-external factor given in (1a). It demonstrates that (32b) is ill-formed because the phi-features of T and the Case feature of the subject can never be valued in the derivation due to the partially moved object in the v*P edge, which blocks phi-feature agreement, with the result that they are transferred unvalued, violating interface conditions and causing crash at the interfaces. We refer the reader to Mizuguchi (2014) for detailed discussion of superiority effects in A′-movement in SMT-based Minimalism, including the discussion of “well-formed” superiority violations.8 Our analysis of A-movement superiority effects in this paper, coupled with Mizuguchi’s (2014) analysis of A′-movement superiority effects, will achieve a unified, extrasyntactic account for superiority effects in light of the interfaces through the transfer of unvalued features, strengthening the Minimalist hypothesis that the properties of language can be explained by free Merge + (1).

REFERENCES

Adger, David (2003) Core Syntax: A Minimalist Approach, Oxford University Press, Oxford.
Bobaljik, Jonathan (1995) Morphosyntax: The Syntax of Verbal Inflection, Doctoral dissertation, MIT.
Bošković, Željko (2002) “A-Movement and the EPP,” Syntax 5, 167–218.
Broekhuis, Hans (2013) “Feature Inheritance versus Extended Projections,” ms., Meertens Institute.
Carstens, Vicki (2011) “Hyperactivity and Hyperagreement in Bantu,” Lingua 121, 721–741.
Castillo, Juan Carlos, John Dury and Kleanthes K. Grohmann (2009) “Merge over Move and the Extended Projection Principle: MOM and the EPP Revisited,” Iberia: An International Journal of Theoretical Linguistics 1, 53–114.
Chomsky, Noam (1986) Knowledge of Language: Its Nature, Origin and Use, Praeger, New York.
Chomsky, Noam (2000) “Minimalist Inquiries: The Framework,” Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik, ed. by Roger Martin, David Michaels and Juan Uriagereka, 89–155, MIT Press, Cambridge, MA.
Chomsky, Noam (2004) “Beyond Explanatory Adequacy,” Structures and Beyond: The Cartography of Syntactic Structures, Volume 3, ed. by Adriana Belletti, 104–131, Oxford University Press, Oxford.

8 Kitahara’s (2011) probe-goal approach to superiority effects in wh-movement also implies that A′-movement superiority effects are reducible to the interfaces.
Chomsky, Noam (2007) “Approaching UG from Below,” Interfaces + Recursion = Language?: Chomsky’s Minimalism and the View from Syntax-Semantics, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29, Mouton de Gruyter, Berlin.

Chomsky, Noam (2008) “On Phases,” Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud, ed. by Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta, 133–166, MIT Press, Cambridge, MA.

Chomsky, Noam (2013) “Problems of Projection,” Lingua 130, 33–49.

Van Craenenbroeck, Jeroen and Marcel den Dikken (2006) “Ellipsis and EPP Repair,” Linguistic Inquiry 37, 653–664.

Doggett, Teal Bissell (2004) All Things Being Unequal: Locality in Movement, Doctoral dissertation, MIT.

Epstein, Samuel D., Hisatsugu Kitahara, Miki Obata and T. Daniel Seely (2013) “Economy of Derivation and Representation,” Cambridge Handbook of Generative Syntax, ed. by Marcel den Dikken, 487–514, Cambridge University Press, Cambridge.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2012) “Structure Building that Can’t Be,” Ways of Structure Building, ed. by Myriam Uribe-Etxebarria and Vidal Valmala, 253–270, Oxford University Press, Oxford.

Epstein, Samuel D. and T. Daniel Seely (2006) Derivations in Minimalism, Cambridge University Press, Cambridge.

Fortuny, Jordi (2008) The Emergence of Order in Syntax, John Benjamins, Amsterdam.

Fukui, Naoki (1986) A Theory of Category Projection and Its Applications, Doctoral dissertation, MIT.

Grohmann, Kleanthes K., John Dury and Juan Carlos Castillo (2000) “No More EPP,” WCCFL 19, 153–166.

Jaeggli, Osvaldo (1986) “Passive,” Linguistic Inquiry 17, 587–622.

Kitahara, Hisatsugu (2011) “Relations in Minimalism,” English Linguistics 28, 1–22.

Kuroda, Shige-Yuki (1988) “Whether We Agree or Not: A Comparative Syntax of English and Japanese,” Linguisticae Investigationes 12, 1–47.

Lasnik, Howard (2003) Minimalist Investigations in Linguistic Theory, Routledge, London.

Marantz, Alec (1984) On the Nature of Grammatical Relations, MIT Press, Cambridge, MA.

Mizuguchi, Manabu (2009) “Extraposition and Cyclic Syntax,” English Linguistics 26, 293–328.

Mizuguchi, Manabu (2013) “Reconsidering Phase-Internal Derivations: Are They Exceptional or Not?” English Linguistics 30, 75–110.

Mizuguchi, Manabu (2014) “Free Merge in Minimalist Syntax: Superiority Effects Reconsidered,” paper presented at the 11th Workshop in General Linguistics.

Narita, Hiroki (2011) Phasing in Full Interpretation, Doctoral dissertation, Harvard University.

Ott, Dennis (2010) “Grammaticality, Interfaces, and UG,” Exploring Crash-Proof Grammars, ed. by Michael T. Putnam, 89–104, John Benjamins, Amsterdam.
Pesetsky, David and Esther Torrego (2007) “The Syntax of Valuation and the Interpretability of Features,” *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, ed. by Simin Karimi, Vida Samiian and Wendy K. Wilkins, 262–294, John Benjamins, Amsterdam.

Richards, Marc (2007) “On Feature Inheritance: An Argument from the Phase Impenetrability Condition,” *Linguistic Inquiry* 38, 563–572.

Sportiche, Dominique (1988) “A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure,” *Linguistic Inquiry* 19, 425–450.

Stowell, Timothy (1981) *Origins of Phrase Structure*, Doctoral dissertation, MIT.

Ura, Hiroyuki (1994) *Varieties of Raising and the Feature-Based Bare Phrase Structure Theory*, *MIT Occasional Papers in Linguistics* 7.

[received March 31, 2014, revised and accepted July 19, 2014]

Division of Language Education
Dokkyo Medical University
880 Kita-Kobayashi, Mibu, Shimotsuga
Tochigi 321–0293
e-mail: mizuguch@dokkyomed.ac.jp