# Agree and Merge repair Labeling

Mariia Privizentseva*

## Abstract

Following the Labeling Algorithm suggested in Chomsky (2013, 2015), movement can apply to repair an otherwise unlabelable syntactic configuration. In this paper, I will develop this idea and suggest that both movement and agreement can be viewed as repair operations for labeling. I will then show that languages with and without EPP differ in whether movement or agreement applies as a basic repair operation in a language.

## 1. Background: Chomsky's Labeling Algorithm

Following Chomsky's more recent work, Merge must be disentangled from Labeling to achieve the simplest possible definition of Merge. Labeling then requires a separate Labeling Algorithm that was suggested in a series of papers by Chomsky (2013, 2015, 2019) and further developed by several authors (see Epstein et al. 2014, 2020, Rizzi 2016, Hayashi 2020, McInnerney 2024, Moro & Roberts 2024 among others). This labeling algorithm constitutes the basis for this paper, so I will lay it out in detail in this section.

This labeling algorithm is set up in a model where Merge is not feature-driven, but applies freely. Consequently, syntactic selectional features that the traditional projection-by-selection approach (see Chomsky 1995, Adger 2003, and also Stabler 1997) relies on are not present. Furthermore, since there is no selection in syntax, labels are factually not used in syntax. Instead, they are required for interpretation at the PF and LF interfaces. As a result, labels can be established later, at Transfer.

The idea underlying Chomsky's labeling algorithm is that labeling applies under minimal search and its outcome depends on the phrase-structural status of merged syntactic objects. Two core configurations are distinguished: Merge

*I am grateful to Fabian Heck and Anke Himmelreich for comments and suggestions. Special thanks to Gereon Müller, whose open-mindedness and theoretical rigorousness made this research possible.

*Strict Cycling: A Festschrift for Gereon Müller*, 407–422
Silke Fischer, Doreen Georgi, Fabian Heck, Johannes Hein, Anke Himmelreich, Andrew Murphy & Philipp Weisser (eds.)
STRICT CYCLING, Universität Leipzig 2024

of a head with a phrase and Merge of two phrases. First, if a head is merged with a phrase, the head, as an atomic computational item, determines the label:[1]

(1) $[_{XP}$ X YP ]

The second relevant configuration is created by Merge of two phrases. Minimal search finds the two heads of the merged phrases and thus does not give an unambiguous result.

(2) $[_{?}$ XP YP ]

There are then two ways to avoid a crash and determine a label. One option is movement. If one of the merged phrases moves out, it thereby turns invisible for the labeling algorithm and the remaining phrase provides a label.

(3) Movement for Labeling
[ YP $[_{ZP}$ Z $[_{XP}$ XP ~~YP~~ ] ] ]

Another option is agreement. If the heads of the two merged phrases agree in some feature, this feature can be taken as a label for the created constituent; see (4). Note that in the latter case, it is not the category that provides a label, but a feature present on both heads. Here and throughout this work I use the notation introduced by Heck & Müller (2007), according to which probe features are indicated as [∗F∗] and Merge/selection features as [•F•].

(4) Agreement for Labeling
$[_{\langle 2\text{PL},\ 2\text{PL}\rangle}$ $[_{XP}$ $X_{[*\phi:\_*]}$ ZP ] $[_{YP}$ $Y_{[\phi:2\text{PL}]}$ WP ] ]

Chomsky (2015) adds a concept of weak heads to this labeling algorithm. Weak heads cannot provide a label by themselves, but must be strengthened by the presence of a specifier. The specifier must remain in its position at Transfer and the label of the final projection is determined via Agree as shown in (5b).

[1] Rizzi (2016) points out that one complication is introduced by head movement: Complex syntactic objects created by head movement must still act as heads despite not being syntactic terminals. This issue has no immediate effect for this paper and I will not focus on it. The solution suggested by Rizzi (2016) is also fully compatible with my proposal.

(5) a. $[_{?}\ \mathrm{X}_{\mathrm{weak}}\ \mathrm{YP}\ ]$
b. $[_{\langle\phi,\phi\rangle}\ \mathrm{ZP}\ [_{\mathrm{XP}}\ \mathrm{X}_{\mathrm{weak}}\ \mathrm{YP}\ ]\ ]$

To sum up, the core idea behind Chomsky's labeling algorithm is that labeling applies at Transfer and a label must be determined simply by minimal search. The actual implementation however includes further non-trivial concepts, most notably the opposition of weak and strong heads (see also Hayashi 2020 for some problems with weak heads) and the idea that otherwise unlabelable configurations may be repaired by movement: It can apply to resolve [XP YP] configurations by moving one of the phrases out and allowing the remaining one to label. Movement also repairs labeling with weak heads: Weak heads are strengthened by movement of some phrase to their specifier. The labeling algorithm thus actively employs the concept of repair, i.e., an operation that would not take place otherwise may apply if the structure would be illegitimate without it. Minimalist syntax does not allow for a straightforward implementation of the repair logic, but it can be naturally accounted for under Optimality Theory (see Grimshaw 1997, Pesetsky 1998, Müller 2000, Heck & Müller 2007, 2013, Heck 2008, 2021 among others).

The goal of this paper is to present a formal implementation of the repairs applying for labeling and more generally operations taking place at Transfer. In doing so, I will focus on the derivation of simple clauses and the puzzling linguistic phenomenon on which Chomsky's approach to labeling offers a novel perspective—EPP. Before turning to my analysis, let's consider the reanalysis of EPP suggested by Chomsky (2013, 2015).

Following Chomsky, the EPP is derived in the following way. At the CP level, there are two projections problematic for labeling. They are indicated as $\alpha$ and $\beta$ in the structure below.

(6) $[_{\mathrm{CP}}\ \mathrm{C}\ [_{\beta}\ \mathrm{T}_{\mathrm{weak}}\ [_{\alpha}\ \mathrm{DP}_{\mathrm{EA}}\ [_{\mathrm{vP}}\ \mathrm{v}\ \ldots\ ]\ ]\ ]\ ]$

Projection $\alpha$ results from Merge of two phrases, so it can be labeled if the heads of the two merged phrases agree or if one of the phrases moves out. Projection $\beta$ is constructed by merging the phrase $\alpha$ and the weak head T. For this projection to get a label, the weak head must be strengthened. Both labeling problems are resolved by movement of $\mathrm{DP}_{\mathrm{EA}}$.

(7) $[_{\mathrm{CP}}\ \mathrm{C}\ [_{\langle\phi,\phi\rangle}\ \mathrm{DP}_{\mathrm{EA}}\ [_{\mathrm{TP}}\ \mathrm{T}_{\mathrm{weak}}\ [_{\mathrm{vP}}\ \sout{\mathrm{DP}_{\mathrm{EA}}}\ [_{\mathrm{v}}\ \mathrm{v}\ \ldots\ ]\ ]\ ]\ ]\ ]$

Movement by assumption renders $DP_{EA}$ invisible in its base position and allows to label $\alpha$ as vP. In its landing site DP strengthens the weak T head, so that it can now label $\beta$. Notably, $DP_{EA}$ must remain in its landing site at Transfer, where labeling applies and consequently the strengthening of T is required. This derives the obligatory EPP in languages like English. Finally, movement of the $DP_{EA}$ also creates a new projection. It is labeled by shared features: The T head must agree with the external argument in $\phi$-features and these features constitute the label.

Languages without EPP differ in that they have a strong, not a weak T head. The strong T can label and does not require strengthening by a specifier.[2] Chomsky however does not address the problem of labeling in a lower vP projection that in the derivation above is also resolved by movement of the external argument. The analysis thus remains incomplete.[3]

In what follows, I will present an analysis of the EPP that builds on Chomsky's labeling algorithm, but differs from the analysis summarized above in that the additional concept of the strength for syntactic heads is not required and the cross-linguistic variation with respect to the EPP follows from which syntactic operation applies as a basic repair operation in a given language. I am thus building up on Chomsky's proposal that internal Merge can apply to repair labeling and suggest that both core syntactic operations—Agree and Merge—are repair operations applying at Transfer. I formalize the analysis using the Harmonic Serialist version of Optimality Theory.

In section 2, I will show how optimization applies at Transfer, present the

[2] Chomsky (2015) further suggests that there is a correspondence between EPP and ECP effects and that the approach captures it. In a nutshell, the ECP also follows from the strength of the T head: Languages with a weak T head require the subject to be present in Spec,TP at Transfer, so that it cannot, for instance, move to Spec,CP and thus remains in the Spell-Out domain. Further extraction is then impossible. This does not happen if the C head is deleted because it then passes its phasehood to the T head. In this case, the subject may remain in Spec,TP and be accessible for operations at the next phase. In languages with a strong T, the requirement for the subject to be in Spec,TP does not occur in the first place.

[3] Assuming that complements of phase heads are rendered invisible in syntax and v is a phase head, one might suggest that once the complement of the v head is transferred, the structure is processed as [$DP_{EA}$ v] and the v head can provide a label as a head (see Narita 2014, as well as Ott 2011, Gallego 2018, and Bayırlı 2022 on labeling and Spell-Out). This solution for labeling in vP is however not fully compatible with Chomsky's original proposal, because the latter assumes that v passes its phasehood to V, so that the specifier of V remains accessible after Spell-Out.

necessary constraints and the analysis of simple clauses. After this, in section 3, I will summarize and discuss some open questions.

## 2. Proposal

### 2.1. Optimization at phase level

Following Chomsky (2008, 2013, 2015, 2019), I assume that after the phase is built by external Merge, it is transferred to the interfaces and the majority of the phase-internal operations, including internal Merge and Agree, apply at Transfer.[4] Merge (external and internal) is not feature-driven and as shown in the previous section internal Merge can apply in order to repair otherwise unlabelable structures. In this paper, I will explore a more radical version of this approach. I would like to propose that Merge as well as Agree are not feature-driven and apply to repair unlabelable structures. I suggest that both operations are restricted by economy constraints, but can take place to avoid a violation of a higher ranked labeling constraint.

I will implement this repair logic (already present in Chomsky's labeling algorithm, but expanded in the current work) and model operations applying at phase level in Optimality Theory (see Sells et al. 1996, Aissen 1997, 1999, 2003, Grimshaw 1997, Müller 1997, Fanselow & Féry 2002, Heck 2008, Müller 2015, Stanton 2016, Murphy 2017 for optimality-theoretic approaches to syntax). In particular, I use Harmonic Serialism, a version of Optimality Theory where the generator is restricted so that the set of outputs is produced by applying at most one operation to the input (see Heck & Müller 2007, McCarthy 2008). Outputs are then evaluated against the constraints and the output with the best constraint profile becomes the input to the next optimization cycle. The derivation terminates when the constraint profile cannot be further improved, that is, the output with the best constraint profile is identical to the input.

[4]For Agree, the necessity to apply at Transfer was motivated by the requirement for features uninterpretable at LF to be deleted before the interface (Chomsky 2004, 2008, Richards 2007). As interpretability of a feature at LF is not visible in syntax, it was suggested that being unvalued in syntax corresponds to being uninterpretable at LF. The valuation (i.e., Agree) must take place at Transfer, because the difference between interpretable and uninterpretable features would be otherwise lost too early. It is however not completely clear whether this logic is applicable to this work as I am assuming that Agree, like Merge, is not driven by probe features; see also Privizentseva (2023) for some problems with this approach.

I would like to suggest that after the phase is built by external Merge, the structure constitutes an input to the Harmonic Serialist derivation where Agree, internal Merge, and labeling apply in the generator and the outputs are evaluated by ranked constraints. I will focus on the derivation of simple clauses and use four constraints. They are introduced in the remainder of this section.

First, presence of a label is one of the requirements for the well-formedness of the syntactic structure and it is ensured by the constraint LABEL:

(8) LABEL:
Every projection present in the input is labeled in the output.

Labeling applies as suggested by Chomsky (2013); that is, labels are determined by minimal search (see also Ke 2024) and there is no distinction between weak and strong heads. This algorithm is summarized in (9). First, all terminal nodes are accessed before the heads of their branching sisters and provide a label. Second, lower copies of displaced syntactic objects are invisible for the search and do not provide labels. Third, if the search finds two heads of merged phrases, it selects a feature they share as a label.

(9) a. X is a label in {X, YP}.
b. XP is a label in {XP, YP} if YP moved out.
c. $\langle \alpha, \alpha \rangle$ is a label in {XP, YP} if XP and YP agree in feature $\alpha$.

Merge and Agree apply freely in the generator. There are no probe features that would require their application, i.e., both Agree and Merge are not feature-driven under this approach. This allows both operations to act as repairs. The outputs are then evaluated against the two economy constraints that restrict the application of Agree and Merge.[5]

(10) *AGREE:
Agree does not apply.

(11) *MERGE:
Merge does not apply.

[5]The *MERGE constraint refers to Merge in general and therefore could in principle restrict the application of external Merge as well, but it will in fact never apply to cases of external Merge, because the constraint is part of optimization that takes place at Transfer, after the structure has already been built by external Merge.

The final constraint that I will call DOMAIN CONSERVATION restricts movement that could in principle apply unboundedly in this approach, since it is not triggered by syntactic features. While the constraint was not formulated as in (12) in previous work, there is a clear precedence for it; see, for instance, the WH-CRITERION by Rizzi (1996), WH-RECOVERABILITY by Heck & Müller (2003). This constraint makes reference to Ā/A-features and Ā/A-positions (cf. Chomsky 1981, Mahajan 1990, Webelhuth 1992, Van Urk 2015, and Keine & Bhatt 2019 among others). I assume that case and $\phi$ are A-features, while [wh], [top], [foc], [rel] are Ā-features. Note that these features, while being present in syntax, do not trigger syntactic operations directly. I am further defining Ā-positions as specifiers of heads that may host Ā-features, but do not necessarily have them in every single derivation. Thus, the complement and the specifier of the V head, the specifier of the v head, and the specifier of the T head are typically A-positions, while the specifier of the C head is the Ā-position.

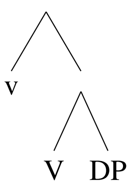
(12) DOMAIN CONSERVATION (DC):
Syntactic objects without Ā-features are in A-positions.

In the next section, I will show how these four constraints derive simple clauses in languages with and without EPP.

## 2.2. Simple clauses and EPP

Let's consider the derivation of a simple clause. Recall that under the current approach, the phase is fully built by external Merge and then the structure is subject to an optimality-theoretic evaluation in the course of which all phase-level operations apply. The structure in (13) constitutes a first input to the Harmonic Serialist optimization procedure.

(13) vP phase



This structure contains two projections that have no labels. They both are created by Merge of a head with a phrase and can be consequently labeled without application of further operations. The first optimization step is shown in (14). The output $O_1$ is identical to the input and it incurs two violations of LABEL, one for each unlabeled projection. In the output $O_2$, one of the projections is labeled in the generator and LABEL is violated only once. Outputs $O_3$ and $O_4$ are built by the application of internal Merge and Agree correspondingly and they violate the respective economy constraints: *MERGE and *AGREE. The ranking of these constraints does not affect the outcome of the optimization at the vP-level, but will become relevant in the derivation of the next phase. Both $O_3$ and $O_4$ have only one violation of LABEL, because the constraint applies only to projections present in the input and in the output. In $O_3$, the DP moves, so that projection [ V DP ] present in the input is not in the output and projection [ DP V ] that is part of the output is not in the input. Consequently, neither of the two projections is visible to the constraint. Similarly, I assume that as agreement changes feature makeup of the head, the projection in the output is not identical to the one in the input and therefore $O_4$ violates LABEL only once.

(14) vP-level optimization, step 1

| | I: [ v [ V DP ] ] | LABEL | DC | *MERGE | *AGREE |
|---|---|---|---|---|---|
| | $O_1$: No changes | ** | | | |
| ☞ | $O_2$: Labeling in VP | * | | | |
| | $O_3$: DP moves to Spec,VP | * | | * | |
| | $O_4$: V agrees with DP | * | | | * |

The second step of the vP-level optimization is shown in (15). Output $O_{24}$ created by labeling of the remaining unlabeled projection has the best constraint profile.[6]

[6] As the optimization is phase-based, labeling in the higher projection could apply before labeling in the lower projection. I assume that this is indeed possible, but the order in which labeling operations apply has no effect on the final output of the derivation.

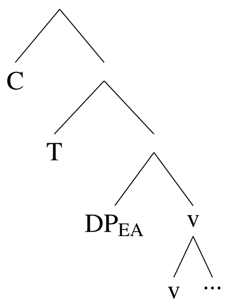
(15) vP-level optimization, step 2

| | $I_2$: [ v [$_{VP}$ V DP ] ] | LABEL | DC | *MERGE | *AGREE |
|---|---|---|---|---|---|
| | $O_{21}$: No changes | * | | | |
| | $O_{22}$: DP moves to Spec,VP | | | * | |
| | $O_{23}$: V agrees with DP | | | | * |
| ☞ | $O_{24}$: Labeling in vP | | | | |

After this, the derivation converges: All projections are labeled; further Agree or Merge do not improve the constraint profile.

I will now turn to the derivation of the CP phase. As in the previous case, the phase built by external Merge (see (16)) is fed into the Harmonic Serialist optimization procedure.

(16) CP phase




The first two steps of optimization at the CP-level are given in (17)-(18). In these steps, projections created by Merge of the head and the phrase are labeled.

(17) CP-level optimization, step 1

| I: | [ C [ T [ DP [$_{vP}$ v ...] ] ] ] | LABEL | DC | *MERGE | *AGREE |
|---|---|---|---|---|---|
| | $O_1$: No changes | *** | | | |
| | $O_2$: DP moves to Spec,TP | ** | | * | |
| | $O_3$: v agrees with DP in $\phi$ | ** | | | * |
| ☞ | $O_4$: Labeling in TP | ** | | | |

(18) CP-level optimization, step 2

| | $I_4$: [C [$_{TP}$ T [ DP [$_{vP}$ v ...] ] ] ] | LABEL | DC | *MERGE | *AGREE |
|---|---|---|---|---|---|
| | $O_{41}$: No changes | ** | | | |
| ☞ | $O_{42}$: Labeling in CP | * | | | |
| | $O_{43}$: v agrees with DP in $\phi$ | * | | | * |
| | $O_{44}$: DP moves to Spec,TP | * | | * | |

The remaining unlabeled projection is created by Merge of the two phrases and cannot be labeled without application of further operations. There are two ways to provide a label: First, the subject DP agrees with *v* and the shared feature is used as a label. Second, the DP moves out and the remaining vP labels. I would like to suggest that both these scenarios are attested, but in different languages. The difference follows from whether Agree or Merge applies as a basic repair operation in a language and resolves the problem created by Merge of the subject. The choice of the repair operation in turn depends on the respective ranking of *MERGE and *AGREE constraints. In languages without EPP, where the subject remains in-situ, *MERGE outranks *AGREE, so that Agree typically repairs labeling.[7]

(19) Languages without EPP
LABEL ≫ DOMAIN CONSERVATION ≫ *MERGE ≫ *AGREE

For these languages, the third optimization step is given in (20). As discussed above, application of Merge and Agree to unlabeled projections remedy one violation of LABEL, because the unlabeled projection in the input is not identical to the unlabeled projection in the output and is therefore disregarded by the constraint. In result, outputs $O_{422}$ and $O_{423}$ have no violations of LABEL. As *MERGE outranks *AGREE, $O_{423}$ formed by agreement wins.

(20) CP-level optimization, step 3

| | $I_{42}$: [$_{CP}$ C [$_{TP}$ T [ DP [$_{vP}$ v ...] ] ] ] | LABEL | DC | *MERGE | *AGREE |
|---|---|---|---|---|---|
| | $O_{421}$: No changes | * | | | |
| | $O_{422}$: DP moves to Spec,TP | | | * | |
| ☞ | $O_{423}$: v agrees with DP in $\phi$ | | | | * |

[7] I assume that the ranking is fixed on the language basis: *MERGE and *AGREE are ranked in all steps of the derivation, but their ranking was irrelevant for the choice of the optimal output in the earlier optimization cycles.

Application of Agree makes labeling by a shared feature possible in the next step (21) and the derivation converges after this.

(21) CP-level optimization, step 4

| $I_{423}$: | [$_{CP}$ C [$_{TP}$ T [ DP [$_{vP}$ $v_{\phi}$ ...] ] ] ] | LABEL | DC | *MERGE | *AGREE |
|---|---|---|---|---|---|
| | $O_{4231}$: No changes | * | | | |
| ☞ | $O_{4232}$: Labeling by $\phi$ | | | | |

Let's now turn to languages with EPP. In these languages, the labeling problem created by the Merge of the external argument is repaired by movement of the subject. Movement is preferred over agreement in situ, because the ranking of constraints banning Merge and Agree is reversed.

(22) Languages with EPP
LABEL ≫ DOMAIN CONSERVATION ≫ *AGREE ≫ *MERGE

The first two optimization steps proceed as shown in (17)–(18) above. The third step is shown in (23) and the output $O_{422}$ where the subject DP moves to Spec,TP has the best constraint profile. Movement to Spec,CP in $O_{424}$ incurs the violation of the DOMAIN CONSERVATION constraint as it places a noun phrase without any Ā-features in the Ā-position.

(23) CP-level optimization, step 3

| $I_{42}$: | [$_{CP}$ C [$_{TP}$ T [ DP [$_{vP}$ v ...] ] ] ] | LABEL | DC | *AGREE | *MERGE |
|---|---|---|---|---|---|
| | $O_{421}$: No changes | * | | | |
| ☞ | $O_{422}$: DP moves to Spec,TP | | | | * |
| | $O_{423}$: v agrees with DP in $\phi$ | | | * | |
| | $O_{424}$: DP moves to Spec,CP | | * | | * |

The [ DP TP ] constituent is created in the third optimization step in (23) and for this reason does not violate the two-level constraint LABEL there. This constituent is part of the input in (24), so it violates LABEL unless further operations apply. There are again two ways to resolve the labeling problem: movement of the DP to a higher projection as in $O_{4222}$ or agreement as in $O_{4223}$. Despite *AGREE being ranked higher than *MERGE, the output $O_{4223}$ with agreement has the best constraint profile. This is because Spec,CP is the only available position for further movement and placement of the subject in this position violates DOMAIN CONSERVATION.[8] As a result, different repairs are preferred at different stages of the derivation within one language.

(24) CP-level optimization, step 4

| $I_{422}$: | $[_{CP}$ C [ DP $[_{TP}$ T $[_{vP}$ v ...] ] ] ] | LABEL | DC | *AGREE | *MERGE |
|---|---|---|---|---|---|
| | $O_{4221}$: No changes | * | | | |
| | $O_{4222}$: DP moves to Spec,CP | | * | | * |
| ☞ | $O_{4223}$: T agrees with DP in $\phi$ | | | * | |

Finally, shared $\phi$-features provide the label in (25).

(25) CP-level optimization, step 5

| $I_{4223}$: | $[_{CP}$ C [ DP $[_{TP}$ $T_{\phi}$ $[_{vP}$ ~~DP~~ $[_{v}$ v ...] ] ] ] | LABEL | DC | *AGREE | *MERGE |
|---|---|---|---|---|---|
| | $O_{42231}$: No changes | * | | | |
| ☞ | $O_{42232}$: Labeling by $\phi$ | | | | |
| | $O_{42234}$: DP moves to Spec,CP | | * | | * |

To sum up, I suggested that after the phase is built by external Merge, the syntactic structure is evaluated against the set of ranked constraints. The constraint LABEL requires the labeling in all syntactic projections, which sometimes forces the application of the core syntactic operations—Agree and Merge. Their application however incurs a violation of the corresponding economy constraints. Languages with EPP and those without it differ in the mutual rankings of the constraints *ARGEE and *MERGE. The possible rankings and the resulting patterns are summarized in the table below:

(26) Languages with and without EPP

| | |
|---|---|
| EPP | *AGREE $\gg$ *MERGE |
| No EPP | *MERGE $\gg$ *AGREE |

## 3. Outlook

This paper makes three contributions. First, investigating Chomsky's labeling algorithm, I show that this algorithm involves a concept of repair, which can be hardly formalized in standard minimalist syntax. The paper presents a formal implementation of Merge and Agree as repair operations in the Harmonic Serialist version of Optimality Theory. Second, I suggest that the attested

[8] If the subject has an Ā-feature, the approach predicts that it may move to Spec,CP directly. This might be a correct prediction (see Messick 2020 among others).

cross-linguistic variation with respect to subject movement follows from whether Merge or Agree applies as a basic repair in a language. Third, this paper explores the possibilities to ensure and correctly restrict the application of Merge and Agree in a system where these operations are not driven by features, but apply freely. I suggested that they are restricted by economy constraints and are acceptable only if their application improves labeling.

The account developed in this paper also raises further questions concerning the analyses of specific linguistic phenomena. I will now identify three such issues.

First, according to this analysis, languages with EPP and without EPP differ in the location of the subject agreement within verbal projections. In languages with EPP, the T head agrees with the subject, while in languages without EPP, the v head is the locus of verbal agreement. On the one hand, this is a falsifiable prediction and indeed different loci of subject agreement were proposed for different languages (see Chomsky 2001, Sigurðsson 2000, Bejar 2003). At the same time, languages with both higher and lower locus of agreement may ultimately display inflection on the verb, for instance, due to post-syntactic morphological operations. They can be thus indistinguishable at the surface.

Second, for languages with EPP, the analysis derives obligatory movement of the argument merged in Spec,vP to Spec,TP, but the analysis so far does not predict movement of unaccusative subjects that originate as complements to the V head. I would like to suggest that subjects of unaccusatives move to Spec,vP because of case filter constraint as in (27) (Chomsky 1981, Vergnaud 2008/1977).

(27) CASE FILTER: All nouns have case.

As subjects of unaccusatives do not receive case at the vP-phase, they must escape the phasal spell-out by movement to the specifier of the phase head, vP, to the position where subjects of unergatives and transitive verbs are first merged. After this, the derivation proceeds as sketched above.

Third, simplifying the model by eliminating the concept of weak heads present in Chomsky's proposal, the approach also narrows its empirical coverage: The analysis does not account for the correspondence between the EPP and the ECP that Chomsky's original proposal is claimed to derive. I suggest that this is a welcome result: The co-occurrence of the two principles is attested in some languages (e.g., in English), but is not universal; see Müller & Rohrbacher (1989).

## References

Adger, David. 2003. *Core syntax: A minimalist approach.* Oxford: Oxford University Press.

Aissen, Judith. 1997. On the syntax of obviation. *Language* 73(4). 705–750.

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17(4). 673–711.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory* 21(3). 435–483.

Bayırlı, İsa Kerem. 2022. Labeling, Concord, and Nominal Syntax in Turkish. *Languages* 7(4). 296.

Bejar, Susana. 2003. *Phi-Syntax: A theory of agreement*: University of Toronto dissertation.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale. A life in language*, 1–52. Cambridge: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *The cartography of syntactic structures: Structures and beyond*, vol. 3, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2008. On Phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In E. Domenico, C. Hamann & S. Matteini (eds.), *Structures, strategies, and beyond: Studies in honour of Adriana Belletti*, 3–16. Amsterdam: John Benjamins.

Chomsky, Noam. 2019. Lectures. Paper presented at University of California, Los Angeles, April 29 – May 2.

Epstein, Samuel D., Hisatsugu Kitahara & T. Daniel Seely. 2020. Unifying labeling under minimal search in “single-" and “multiple-specifier" configurations. In Samuel D. Epstein, Hisatsugu Kitahara & T. Daniel Seely (eds.), *A minimalist theory of simplest merge*, 140–148. New York, London: Routledge.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2014. Labeling by Minimal Search: Implications for successive-cyclic A-movement and the conception of the postulate "phase". *Linguistic Inquiry* 45(3). 463–481.

Fanselow, Gisbert & Caroline Féry. 2002. *Resolving conflicts in grammars: Optimality Theory in syntax, morphology and phonology*. Hamburg: Buske.

Gallego, Ángel J. 2018. Projection without agreement. *The Linguistic Review* 35(4). 601–623.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28(3). 373–422.

Hayashi, Norimasa. 2020. Labeling without weak heads. *Syntax* 23. 275–294.

Heck, Fabian. 2008. *On pied-piping – wh-movement and beyond*. Berlin: Mouton de Gruyter.

Heck, Fabian. 2021. Syntactic repairs and cyclic optimization. Ms., Leipzig University.

Heck, Fabian & Gereon Müller. 2003. Derivational optimization of Wh-movement. *Linguistic Analysis* 33(1-2). 97–148.

Heck, Fabian & Gereon Müller. 2007. Extremely local optimization. In Erin Brainbridge & Brian Agbayani (eds.), *WECOL 34: Proceedings of the 34th Western Conference on Linguistics*, 170–183. California State University: Fresno.

Heck, Fabian & Gereon Müller. 2013. Extremely local optimization. In Ralf Vogel & Hans Broekhuis (eds.), *Linguistic derivations and filtering*, 136–165. Sheffield: Equinox Publishing.

Ke, Alan Hezao. 2024. Can Agree and Labeling Be Reduced to Minimal Search? *Linguistic Inquiry* 55(4). 849–870.

Keine, Stefan & Rajesh Bhatt. 2019. Secondary strong crossover in Hindi and the typology of movement. In Maggie Baird & Jonathan Pesetsky (eds.), *NELS 49: Proceedings of the Forty-Ninth Annual Meeting of the North East Linguistic Society*, 125–134. Amherst, MA: GLSA.

Mahajan, Anoop. 1990. *The A/A-bar distinction and movement theory*. Cambridge, MA: MIT dissertation.

McCarthy, John. 2008. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26. 499–546.

McInnerney, Andrew. 2024. The Position of Wh-Subjects in Labeling Theory. *Linguistic Inquiry* 55(3). 579–594.

Messick, Troy. 2020. The derivation of highest subject questions and the nature of the EPP. *Glossa: A journal of general linguistics* 5(1). 13.

Moro, Andrea & Ian Roberts. 2024. The duality of syntax: Unstable structures, labelling and linearisation. *Natural Language & Linguistic Theory* 42. 609–631.

Müller, Gereon. 1997. Partial Wh-movement and Optimality Theory. *The Linguistic Review* 14(3). 249–306.

Müller, Gereon. 2000. Das Pronominaladverb als Reparaturphänomen. *Linguistische Berichte* 182. 139–178.

Müller, Gereon. 2015. Optimality-theoretic syntax. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and Analysis. An International Handbook*, 875–936. Berlin: Mouton de Gruyter.

Müller, Gereon & Bernhard Rohrbacher. 1989. Eine Geschichte ohne Subjekt. Zur Entwicklung der pro-Theorie. *Linguistische Berichte* 119. 3–52.

Murphy, Andrew. 2017. *Cumulativity in syntactic derivations*. Leipzig: Leipzig University dissertation.

Narita, Hiroki. 2014. *Endocentric structuring of projection-free syntax*. Amsterdam: John Benjamins.

Ott, Dennis. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42(1). 183–192.

Pesetsky, David. 1998. Optimality principles of sentence pronunciation. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 337–383. Cambridge, MA: MIT Press.

Privizentseva, Mariia. 2023. Nominal ellipsis reveals concord in Moksha Mordvin. *Syntax* 26. 355–403.

Richards, Marc. 2007. On Feature Inheritance: An Argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38(3). 563–572.

Rizzi, Luigi. 1996. Residual verb second and the Wh-Criterion. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and functional heads: Essays in comparative syntax*, vol. 2, 63–90. Oxford: Oxford University Press.

Rizzi, Luigi. 2016. Labeling, maximality and the head – phrase distinction. *The Linguistic Review* 33(1). 103–127.

Sells, Peter, John Rickford & Thomas Wasow. 1996. An optimality theoretic approach to variation in negative inversion in AAVE. *Natural Language and Linguistic Theory* 14(3). 591–627.

Sigurðsson, Halldor Armann. 2000. The locus of case and agreement. *Working Papers in Scandinavian Syntax* 65. 591–627.

Stabler, Edward P. 1997. Derivational minimalism. In Christian Retoré (ed.), *Logical aspects of computational linguistics*, 68–95. Berlin: Springer.

Stanton, Juliet. 2016. Wholesale late merger in Ā-movement: Evidence from preposition stranding. *Linguistic Inquiry* 47(1). 89–126.

van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor*. Cambridge, MA: MIT dissertation.

Vergnaud, Jean-Roger. 2008/1977. Letter to Noam Chomsky and Howard Lasnik on "Filters and control," April 17, 1977. In Carlos P. Otero Robert Freidin & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 3–15. Cambridge, MA: MIT Press.

Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. Oxford: Oxford University Press.