1. Introduction

Movement is one of the fundamental properties of human language. Within the framework of the Principles and Parameters Theory (see Chomsky (1981, 1986) and Chomsky and Lasnik (1993)), this property is captured by Move-alpha, in which an element may move to a different position to meet some grammatical requirement, leaving a coindexed trace, which is a phonetically-null element that has the relevant properties of the moved element, such as its theta-role. The moved element and its trace(s) also form a (nontrivial) chain.

Within the Minimalist Program (see Chomsky (1993, 1994, 1995, 2000, 2001)), the properties of movement, traces, and chains are reevaluated in light of virtual conceptual necessity. In particular, only the interface levels LF and PF are assumed, and the addition of new objects in the course of the derivation from the numeration to LF is banned (the Inclusiveness Condition). Addressing the issues of reconstruction, Chomsky (1993) proposes the “copy theory of movement”: a trace is a copy of the moved element, deleted in the phonological component in

* I would like to thank Hirohisa Kiguchi, Mitsue Motomura, and Shigeki Taguchi and two anonymous reviewers for invaluable comments and criticisms on earlier version of this paper. I am also indebted to Charlie Robertson for his suggestions for improving the present paper. All errors are of course mine. The research presented here is in part supported by Grants-in-Aid for Young Scientists (B) (No. 18720113), the Ministry of Education, Culture, Sports, Science, and Technology.
the case of overt movement. This theory has an advantage in that it is compatible with the Inclusiveness Condition, and it dispenses with the reconstruction operation from the grammar. However, as Nunes (2004: 2) correctly points out, "if no explanation is provided for why 'traces' (lower copies) must be deleted in the phonological component, the notion of trace as a primitive is being reintroduced."

Given the copy theory of movement, the important question that arises is why the traces cannot be phonetically realized rather than the head of a chain. Nunes (2004) tries to answer this question, and argues that all members of chain links are subject to the same principle, and that the phonetic realization of a copy in a nontrivial chain is determined by properties of the computational system. More specifically, deletion of copies is forced; otherwise, no linear order can be established in accordance with Kayne's (1994) Linear Correspondence Axiom (LCA).

The organization of this paper is as follows. In Section 2, I first see Nunes's (2004) mechanism of how a member of a given chain is phonetically realized. In particular, the syntactic operations Chain Reduction and Formal Feature Elimination (FF-Elimination) are discussed. In Section 3, I will observe that sideward movement, as a consequence of the Copy + Merge theory of movement, can derive some of the properties of parasitic gap constructions. In Section 4, I will point out two problems that arise from Nunes's (2004) analysis of the linearization of chains and I will try to make some revisions of his analysis. Finally, Section 5 contains concluding remarks.

2. Linearization and the LCA

2.1. The Questions

Chomsky (1993) revives the copy theory of movement, according to which, in the case of overt movement, a moved element leaves a copy that is deleted in the phonological component but remains available for interpretation at LF. One of the advantages of this theory is that it satisfies the Inclusiveness Condition, which requires that an LF object be built from the features of the lexical items of the numeration (see Chomsky (1995: 228)). Note that a trace is a violation of the Inclusiveness Condition, since it is introduced in the course of the derivation.

Given the copy theory of movement, Nunes (2004: 17) addresses the following questions: (i) why is it the case that (in general) a nontrivial
SOME REMARKS ON LINEARIZATION OF CHAINS

chain cannot have all of its links phonetically realized?, and (ii) why is it the case that (in general) traces and not heads of chains are the links that are deleted? More specifically, given the derivation in (1), why is it that the structure in (1c) cannot yield the PF output related to (2a) and (2b)?

(1)  
   a. \( K = [_{TP} \ T \ [_{VP} \ was \ [_{VP} \ kissed \ John]]] \)  
   b. Copy  
      \( K = [_{TP} \ T \ [_{VP} \ was \ [_{VP} \ kissed \ John'i]]] \)  
      \( L = John'i \)  
   c. Merge  
      \([_{TP} \ John'i \ [_{\Gamma} \ T \ [_{VP} \ was \ [_{VP} \ kissed \ John'i]]] \)  

(2)  
   a. *John was kissed John.  
   c. John was kissed.

The issues become more complex when we consider the following data, where the copies can be phonetically realized in addition to the head of the chain.

(3)  
German (from McDaniel (1986))  
   *Wen glaubt Hans wen Jakob gesehen hat?  
   Whom thinks Hans whom Jakob seen has  
   'Who does Hans think Jakob saw?'

Thus, we need to seek an analysis that permits copies to be phonetically realized in some cases but restricted as to not allow for pronunciation of every copy. Nunes (2004) argues that the reason why (in general) traces are not phonetically realized follows from the LCA and economy considerations, which will be observed below.

2.2. Nondistinctness

The important question under the copy theory of movement is whether the two terms with the same set of features are interpreted as distinct elements or copies. Let us now consider (4). In (4), John in the object position raises to TP SPEC and it leaves behind a copy. If two occurrences of John are elements distinctively selected in the numeration, then no chain is formed, and it does not obey the LCA for linearization purposes.

(4)  
   \([John'i \ [was \ [kissed \ John'i]]] \)

Nunes (2004) simply follows Chomsky (1995), who proposes that two lexical items 1 and 1' selected from a numeration should be marked as distinct for the computational system if they are accessed by distinct
applications of Select.

The question that arises is whether the term created by the Copy operation is identical rather than nondistinct. Nunes (2004: 164, fn. 14) answers this question stating that "the motivation for treating terms related by the Copy operation as nondistinct rather than identical has to do with feature checking." For instance, the two copies of *John* in (4) are not identical, since the upper copy of *John* checks its Case-feature against T, whereas the Case feature of the lower copy remains unchecked. Therefore, under Nunes's (2004) framework, in a chain \((\alpha_1, \alpha_2)\), \(\alpha_1\) is not identical to \(\alpha_2\), but is nondistinct from \(\alpha_2\), in terms of its feature composition of the terms.

2.3. LCA

The main point of Kayne's (1994) influential work is that linear order is determined by hierarchical structure, as stated by the Linear Correspondence Axiom (LCA).

(5) Linear Correspondence Axiom

Let \(X, Y\) be nonterminals and \(x, y\) terminals such that \(X\) dominates \(x\) and \(Y\) dominates \(y\). Then if \(X\) asymmetrically c-commands \(Y\), \(x\) precedes \(y\).

The question that arises is whether a copy of the moved element is computed for purpose of linearization, assuming that the links of a chain are (in a sense) the same element. Given the copy theory of movement, the null hypothesis should dictate that all copies obey the same grammatical conditions. Nunes (2004) takes this issue seriously, and he concludes that if two links of a given chain are formed by a movement operation, any material intervening them asymmetrically c-commands and is asymmetrically c-commanded by the same element. Let us consider the structure in (6) in order to determine how the structure should be linearized in accordance with the LCA.

(6) \([John^i \quad [was \quad [kissed \quad John^i]]]\)

Let us focus on the relationship between the two copies of *John* and the copula *was*. The upper copy of *John* asymmetrically c-commands *was*, which means that the order \(<John^i, \ was>\) is obtained, according to the LCA.\(^1\) However, the order \(<\ was, \ John^i>\) is gained since the copula

---

\(^1\) It is not clear to me whether the upper copy of *John* asymmetrically c-commands the copula, if we follow the structure in (1c) (see also Gärtner (1998)).
asymmetrically c-commands the lower copy of John. If we combine the results, we obtain the sequence \( \langle \text{John}^i, \text{was}, \text{John}^i \rangle \). Since the two instances of John are nondistinct, this sequence lacks asymmetry and the copula must precede and be preceded by the same element John. This sequence also violates the irreflexivity condition on linear order (if \( \alpha \) precedes \( \beta \), then it must be the case that \( \alpha \neq \beta \)).

Nunes (2004) proposes that in order to avoid an undesirable result, the deletion of chain links is required for a structure with nontrivial chains to be linearized in accordance with the LCA. Thus, one of the copies in (6) must be deleted in order to avoid a violation of the asymmetry and irreflexivity conditions on linear order.

2.4. Chain Reduction

If we propose that the deletion of chain links is required in order not to violate the linearization conditions in accordance with the LCA, the question that immediately arises is how the chain links are deleted. In order to see this, let us first consider the following structure:

(7) \[ \text{TP [DP the [NP tall man]]}^i \text{ appears [TP [DP the [NP tall man]]}^i \text{ to have been kissed [DP the [NP tall man]]}^i] \]

Given that these three copies are nondistinct, this structure cannot be linearized as it is, since the highest copy of the tall man asymmetrically c-commands the verb appears, which asymmetrically c-commands the lower copies of the tall man. Therefore, the deletion of chain links must take place in order to linearize the structure without inducing the asymmetry and irreflexivity conditions in accordance with the LCA.

Nunes (2004: 27) proposes that the deletion of chain links is subject

According to the LCA and the definition of c-command (i) proposed by Kayne (1994), the copula and the upper copy of John c-commands the other, so that the linear order between John and was cannot be determined.

(i) c-command

\( X \) c-commands \( Y \) iff \( X \) and \( Y \) are categories and \( X \) excludes \( Y \) and every category that dominates \( X \) dominates \( Y \). (Kayne (1994: 16))

In order to exclude the problem, the structure in (1c) would be reanalyzed and the subject NP should appear in the TP adjoined position.

(ii) \[ \text{TP John}^i \text{ TP T [VP was [VP kissed John}^i]]] \]

It seems that the whole argument can still be maintained even though the subject is located in the TP adjoined position. Thus, I will continue to follow Nunes (2004), where the subject undergoes movement to TP SPEC, not to the TP adjoined position.
to the condition called *Chain Reduction*, as stated below:

(8) *Chain Reduction*

Delete the minimal number of constituents of a nontrivial chain CH that suffices for CH to be mapped into a linear order in accordance with the LCA.\(^2\)

Chain Reduction can correctly exclude the deletion within the different links of a chain called scatted deletion, as shown in (9):

(9) \[TP [DP the [NP tall man]] \text{ appears } [TP [DP the [NP tall man]] \text{ to have been kissed } [DP the [NP tall man]]] \]

Note that the structure in (9) does not violate the asymmetry and irreflexivity conditions on linear order. However, the derived sentence in (10) is unacceptable.

(10) *The appears tall to have been kissed man.*

The reason why (10) is not acceptable is due to the fact that the deletion applies (at least) five times to derive the sentence in (10), which is not the most economical derivation starting from (7). Note that the following structures derived from (7) are more economical, because deletion applies only twice. Furthermore, the structures in (11), yielding the sentences in (12), can be linearized in accordance with the LCA; and thus, Chain Reduction correctly blocks the structure in (9), which requires (at least) five applications of the deletion process.

(11) a. \([TP [DP the [NP tall man]] \text{ appears } [TP [DP the [NP tall man]] \text{ to have been kissed } [DP the [NP tall man]]] \]

b. \([TP [DP the [NP tall man]] \text{ appears } [TP [DP the [NP tall man]] \text{ to have been kissed } [DP the [NP tall man]]] \]

c. \([TP [DP the [NP tall man]] \text{ appears } [TP [DP the [NP tall man]] \text{ to have been kissed } [DP the [NP tall man]]] \]

(12) a. The tall man appears to have been kissed.

b. *Appears the tall man to have been kissed.

c. *Appears to have been kissed the tall man.

Chain Reduction predicts that if each of the three links of the DP chain in (7) has been deleted, the resulting structure in (13) would be ungrammatical, even though the asymmetry and irreflexivity conditions

\(^2\) As Nunes (2004) explicitly notes, it is not necessary to specify that Chain Reduction deletes the minimal number of constituents. Rather, economy conditions concerning the length of derivations can determine the number of elements (copies) to be deleted.
SOME REMARKS ON LINEARIZATION OF CHAINS

on linear order have been avoided.

(13) [TP [DP [NP [tall man]]] appears [TP [DP [NP [tall man]]] to have been kissed [DP [NP [tall man]]]]

(14) *Appears to have been kissed.

In (13), Chain Reduction has been applied three times to the structure. Note, however, that only two applications of deletion are sufficient for the structure containing the DP chain to be mapped into a linear order, as illustrated in (11). Therefore, the structure in (13) is not economical, and it would be ruled out, as desired.³

In this way, recoverability of deletion (in the case of Chain Reduction) can be guaranteed because of the economy conditions regarding the length of derivations. In other words, Chain Reduction ensures that (at least) one copy of chain links must be phonetically realized.

2.5. Formal Feature Elimination

Given the LCA, some of the copies must be deleted in order to satisfy the asymmetry and irreflexivity conditions on linear order. The next question that needs to be answered is which copy is phonetically realized. Let us now consider the structure in (6), for instance. The deletion of either copy and the deletion of both copies satisfy the LCA. However, the structure in (15a) is unacceptable.

   b. John was kissed.

The question is how (15a) is excluded. Following Chomsky’s (1995) proposal concerning the deletion of formal features in a derivation, Nunes (2004) argues that the deletion of chain links in accordance with the LCA is connected to the elimination of formal features in the phonological component. Formal features are divided into two types, interpretable and uninterpretable features. Chomsky (1995) proposes that the deletion of uninterpretable features takes place under feature checking, where deletion makes a given feature “invisible at LF but accessible to the computation” (Chomsky (1995: 280)). Because interpretable features need not be deleted, they are always accessible to the

³ Nunes (2004) argues that even though it is not economical, scattered deletion may yield a convergent derivation. See Nunes (2004: 25–30) for detailed discussion.
computational system, which means that they can establish more than one checking relation. On the other hand, uninterpretable features need to be deleted, and thus they cannot enter into a checking relation once checked. Extending Chomsky’s (1995) checking theory, Nunes (2004) assumes that checking operations render the uninterpretable features invisible at PF and LF.

Moreover, it appears that formal features are relevant for morphological computations in the phonological component. However, formal features are not legible elements at the PF level, and thus there must be an operation of the phonological component applying after Morphology that eliminates the formal features which are visible at PF (see Chomsky (1995)). Nunes (2004) calls this operation Formal Feature-Elimination (see also Nunes (1995)):

\[(16) \text{ Formal Feature Elimination (FF-Elimination)}^{4}\]

Given the sequence of pairs \(\sigma = \langle (F, P)_1, (F, P)_2, \ldots, (F, P)_n \rangle\) such that \(\sigma\) is the output of Linearize, \(F\) is a set of formal features, and \(P\) is a set of phonological features, delete the minimal number of features of each set of formal features in order for \(\sigma\) to satisfy Full Interpretation at PF.

Given FF-Elimination, we have an account of why the head of a chain is generally phonetically realized. Let us reconsider the structure (6), which is repeated below for convenience.

\[(17) \text{[John}_i \text{ [was [kissed John}_i\text{]]]}\]

The upper copy of John raises to TP SPEC in order to check strong D-features, and the Case-feature of John is also deleted. Since Case is an uninterpretable feature, this feature checking renders the feature invisible at LF and at PF. On the other hand, the Case feature of the lower copy remains unchecked.

Let us now examine how pronunciation of copy links is determined. The possible outputs of (6) are illustrated below:5

\[(18) \begin{align*}
\text{a. [John}_i\text{-CASE [was [kissed John}_i\text{-CASE]]]} \\
\text{b. [John}_i\text{-CASE [was [kissed John}_i\text{-CASE]]]} 
\end{align*}\]

---

4 As in the case of Chain Reduction, Nunes (2004: 166, fn. 21) notes that the economy considerations regarding derivational length can specify the number of features to be deleted. Thus, a derivation in which a given feature \(F\) is unnecessarily deleted by FF-Elimination is derivationally longer than a derivation where \(F\) is not deleted.

5 Unchecked features are in bold and checked features are subscribed.
Both (18a, b) satisfy Chain Reduction. The output in (18b) is optimal because no application of FF-Elimination is necessary. However, (18a) is not an optimal output because FF-Elimination is required to delete the Case-feature of the lower copy in order for Full Interpretation to be satisfied. Therefore, the upper copy of *John* is phonetically realized, as desired.

2.6. Multiple Copies

So far, we have observed that Chain Reduction coupled with FF-elimination correctly determines which copy of a given nontrivial chain is phonetically realized. However, as shown below, there are cases where two or more copies of a given chain are phonetically realized. In this subsection, we will examine Nunes’ (2004) argument, which illustrates one of the motivations why the nontrivial chains must be subject to the asymmetry and irreflexivity conditions in accordance with the LCA.

(19) German (from McDaniel (1986))

> Wen glaubt Hans wen Jakob gesehen hat?
> Who thinks Hans whom Jakob seen has
> ‘Who does Hans think Jakob saw?’

In order to account for the phonetic realization of multiple copies in a given nontrivial chain, Nunes (2004) relies on Chomsky (1995) regarding linearization of two heads in a relation of mutual c-command, as illustrated below:

(20)

```
  L
 / \
 m   p
```

In this scenario, neither *m* nor *p* in (20) asymmetrically c-commands the other and no linear order between them can be determined in accordance with the LCA. In order to avoid this problem, Chomsky (1995: 337) suggests that “the structure \( N = [L, m, p] \) has changed by the time the LCA applies so that its internal structure is irrelevant; perhaps \( N \) is converted by Morphology to a ‘phonological word’ not subject internally to the LCA, assuming that the LCA is an operation that applies after Morphology.”

Given this line of reasoning, Nunes (2004) argues that the example

---

6 See Kayne (1994: 38–41) for the possibility of applying the LCA below the word level. See also Nunes (2004: 168, fn. 33) for the empirical evidence that the LCA cannot be applied word-internally.
(19) can be explained in the following way: First, the \(wh\)-phrases in these languages adjoin to an intermediate \(C^0\). Second, Morphology in these languages may convert the adjunction structure \([C^0 \text{WH} [C^0 C^0]]\) into a single terminal element, following Chomsky (1995). The idea here is that the copy in \(C^0\) is not subject to the LCA because the copy and \(C^0\) can count as a single word due to morphological reanalysis (morphological fusion in the sense of Distributed Morphology). That is why Chain Reduction cannot operate the copy in \(C^0\); and thus, the intermediate \(wh\)-copy may be phonetically realized.

Since morphological fusion is concerned with heads, not with maximal projections (full phrases), the analysis above predicts that the intermediate copy must be deleted if \(wh\)-movement involves full phrases, such as \(whose\) \(book\). The unacceptable example below shows that this prediction is borne out:

(21) German

\[*Wessen Buch glaubst du wessen Buch Hans leist?*

\[whose\] \[book\] \[think\] \[you\] \[whose\] \[book\] \[Hans\] \[reads\]

\(Whose\) \(book\) \(do\) \(you\) \(think\) \(Hans\) \(is\) \(reading?\)"

The \(wh\)-phrase in (21) is a full phrase, and thus it fails to adjoin to \(C^0\), and morphological fusion never occurs. This means two things: first, the intermediate copy is subject to the LCA, and second, Chain Reduction dictates that it must be deleted, due to the FF-Elimination.

The discussion above suggests that the pronunciation of more than one copy happens only if the LCA is inapplicable to the structure, due to morphological fusion. To put it another way, the links of a given nontrivial chain must be subject to the LCA.

2.7. Summary

In this section, I have summarized the main argument of Nunes (2004) concerning the linearization of nontrivial chains, assuming the null hypothesis that every member of chain links is subject to the same grammatical operations. The interaction of the LCA with economy considerations (such as FF-Elimination) determines which term of a nontrivial chain is phonetically realized. One member of a chain must

---

7 It is not the case that all languages allow morphological reanalysis, which means that there is a variation among languages depending on whether morphological reanalysis is permitted.
be pronounced because of the operation Chain Reduction, which (generally) deletes all but one link of a nontrivial chain.

3. A New Theory of Movement


Within the framework of Principles and Parameters theory (Chomsky (1981, 1986), Chomsky and Lasnik (1993)), D-structure provides phrasal syntactic objects. However, under the framework of the Minimalist Program (Chomsky (1993, 1994, 1995)), D-structure should be eliminated due to virtual conceptual necessity, and thus the computational system needs to provide means to assemble phrasal syntactic objects. Reintroducing generalized transformation, Chomsky (1994, 1995) proposes that Merge is a building operation that takes two syntactic objects $\alpha$ and $\beta$ and forms the complex syntactic object $K = \{ \gamma \{ \alpha, \beta \} \}$, (substitution) or $L = \{ \langle \gamma, \gamma \rangle, \{ \alpha, \beta \} \}$ (adjunction), where $\gamma$ indicates the label of the resulting object.

Chomsky (1994, 1995) further proposes that the operation Move should be reformulated in the following way: given the syntactic object $\Sigma$ with constituent $K$ and $\alpha$, Move targets $K$, raises $\alpha$, and merges $\alpha$ with $K$, forming $\Sigma'$; operation is cyclic if $\Sigma = K$ and noncyclic otherwise. Move also forms the chain $CH = (\alpha, t)$, where $t$ (the trace of $\alpha$) is a copy of $\alpha$ that is deleted in the phonological component in the case of overt movement. But the deleted copy remains available for interpretation at LF. More simply put, the operation Move can be decomposed into four suboperations, as described below:

(22) Suboperations of Move
a. Copy
b. Merge
c. Form Chain
d. Delete Traces for purposes

3.2. Copy + Merge Theory of Movement

Nunes (2004) examines the suboperations of Move, and points out the problems with (22b), (22c) and (22d). Let us first consider Merge. Merge is conceptually motivated; otherwise, no syntactic objects can be created. However, Nunes (1994) points out that it is strange that under Chomsky's (1994, 1995) system, Merge is an independent operation in certain cases, and at the same time, it is a suboperation of Move in
other cases (see also Gärtner (1997)).

The operation Form Chain is also problematic. Since chain formation and Move expresses the same type of relation, Form Chain would be redundant in the computational system (see also Brody (1995)).

Finally, as observed in the previous section, the deletion of copies reflects optimal applications of the independent operation Chain Reduction, which is needed for linearization purposes. Under Nunes’s analysis, trace is not a grammatical primitive object, and all links of a chain share the same status in that they are subject to the same syntactic operations. This analysis clearly shows that there cannot be a suboperation like (22d), which simply deletes traces.

Addressing these concerns, Nunes (2004) proposes that Move should be decomposed into independent operations, as described below. This is what he refers to as the Copy + Merge theory of movement.

(23) Independent operations of the Copy + Merge theory of movement
a. Copy
b. Merge
c. Form Chain
d. Chain Reduction

Under this approach, the conceptual problems with the operation Move in Chomsky (1994, 1995) can be overcome. Merge is not a suboperation of Move since Move is not an independent operation under the Copy + Merge theory of movement. Second, by dissociating the operations Copy and Merge from Form Chain, the redundancy between Form Chain and Move, criticized by Brody (1995), is eliminated. In other words, as discussed below, Copy + Merge theory of movement allows certain instances where movement does not create a nontrivial chain. Finally, Chain Reduction dispenses with the problematic operation in (22d).

One of the consequences of the Copy + Merge theory of movement is that it allows for what Nunes (1995, 2001, 2004) calls sideward movement. Note that if movement is the interaction of independent operations, such as Copy and Merge, the computational system can cre-

---

8 Sideward movement shares the main characteristics with Interarboreal Operations proposed by Bobaljik and Brown (1997) and Paracyclic movement proposed by Uriagereka (1998).
ate a copy $\alpha$ of an element from one tree and merge $\alpha$ to another tree, as schematized below:

\begin{enumerate}
\item \[ \ldots \alpha \ldots \] \[ \ldots \beta \ldots \] \ (two independent trees)
\item \[ \ldots \alpha \ldots \] $\alpha$ \[ \ldots \beta \ldots \] \ (the copy of $\alpha$)
\item \[ \ldots \alpha \ldots \] \[ \alpha \ldots \beta \ldots \] \ (Merging the copy of $\alpha$ with the other tree containing $\beta$)
\end{enumerate}

Notice that in the case of sideward movement, the landing site of the moved element does not c-command its launching site.

Nunes (2004) argues that sideward movement can derive the core properties of parasitic gap constructions (as well as across-the-board extraction constructions). Before examining a detailed analysis of parasitic gap constructions, I will outline conditions on chain formation in the next subsection.

3.3. Form Chain

Under the analysis of Nunes (2004), the operation Form Chain plays an important role in overt syntax as well as at PF. The main reason for this is that Chain Reduction must apply to a nontrivial chain; otherwise, the resultant chain would violate the asymmetry and irreflexivity conditions on linear order. Furthermore, the Copy + Merge theory of movement predicts that it is possible that an element can move to a position that does not c-command it. This does not mean, contrary to Chomsky (1994, 1995), that movement always creates a nontrivial chain. As given in (25)-(27), Nunes (2004: 91) states the conditions of Form Chain and the auxiliary definitions of sublabel and closeness.

\begin{enumerate}
\item \textit{Conditions on Form Chain}
\item Two constituents $\alpha$ and $\beta$ can form the nontrivial chain $\text{CH} = (\alpha, \beta)$ if
\item a. $\alpha$ is nondistinct from $\beta$;
\item b. $\alpha$ c-commands $\beta$;
\item c. there is at least one feature $F$ of $\alpha$ such that $F$ enters into a checking relation with a sublabel of the head of the projection with which $\alpha$ merges and for any such
\end{enumerate}

In recent years, the sideward movement analysis has been extended to obligatory control (Hornstein (2001), PRO-gate effects (Kiguchi and Hornstein (2004)), donkey anaphora (Boeckx (2003)), extraction domains (Nunes and Uriagereka (2000)), and gapping and pseudo-gapping (Agbayani and Zoerner (2004)), and many others.
feature F of α, the corresponding feature F of β is accessible to the computational system; and
d. there is no constituent γ such that γ has a feature F' that is of the same type as the feature F of α, and γ is closer to α than β is.

(26) **Sublabel**
σ is a sublabel of the head H iff σ is a feature of H or a feature of some element adjoined to H.

(27) **Closeness**
γ is closer to α than β is iff (a) α c-commands γ and γ c-commands β and (b) γ is not in the same minimal domain as α or β.

(25a) indicates that a chain can only be formed between copies of a given syntactic object. Nunes (2004) introduces Chomsky's (1995) Last Resort (25c) and Minimal Link Condition (25d) in the conditions on Form Chain.

3.4. Parasitic Gap Constructions

In this section, I will examine how parasitic gap constructions exemplified in (28) can be derived in the Copy + Merge theory of movement.

(28) Which paper did you file without reading?
Suppose that the computational system starts with the initial numeration N in (29) and creates the derivational stage in (30), where N has been reduced to N' and the two syntactic objects K and L have been assembled.

(29) N = {which₁, paper₁, Q₁, you₁, did₁, v₂, file₁, without₁, C₁, PRO₁, T₁, reading₁}

(30) a. N' = {which₀, paper₀, Q₁, you₁, did₁, v₁, file₀, without₁, C₀, PRO₀, T₀, reading₀}
   b. K = [[CP C [[[TP PRO] [T T [[vP t₁ [v [VP reading [which paper]]]]]]]]]
   c. L = file

At this point, the verb file must merge with the two elements that could bear an object theta-role and the subject theta-role. Unfortunately, the derivation does not converge, since there is only one possible candidate remaining in N' that could have a theta-role, namely you. The derivational step outlined in (30), however, can produce a convergent syntactic object if the computational system makes a copy of which paper in the
adjunct clause and merges it with L as an instance of sideward movement, as illustrated in (31):

\[(31)\quad a. \quad K = [\text{CP } C [\text{TP } \text{PRO}] [\text{T} T [\text{vP } t_j [\text{v} \text{v} [\text{vP reading [which paper]i]]]\\]
\quad b. \quad M = [\text{vP file [which paper]i}]\\
\]

If we assume that movement into theta positions is a licit operation (see Bošković (1994), Lasnik (1995), Bošković and Takahashi (1998), Hornstein (1999, 2001), among others), this sideward movement is also licit because it satisfies the thematic properties of the verb *file*.

Now, as the derivation proceeds and other lexical items in N’ in (30a) merge with K and M in (31), we have two syntactic objects P and Q, as shown in (32):

\[(32)\quad a. \quad P = [\text{PP without } [\text{CP } C [\text{TP } \text{PRO}] [\text{T} T [\text{vP } t_j [\text{v} \text{v} [\text{vP reading [which paper]i]]]\\]
\quad b. \quad Q = [\text{vP you [v} \text{v} [\text{vP file [which paper]i]]]]\\
\]

Then, P adjoins to Q in the next derivational step. Notice that the copies of *which paper* do not form a nontrivial chain, since they are not in a c-command relation.

\[(33)\quad [\text{vP } [\text{vP you [v} \text{v} [\text{vP file [which paper]i]]] [\text{PP without } [\text{CP } C [\text{TP } \text{PRO}] [\text{T} T [\text{vP } t_j [\text{v} \text{v} [\text{vP reading [which paper]i]]]\\]
\]

After the remaining lexical items of N’ merge with the structure in (33) and *you* and *did* move, the structure in (34) is obtained. Finally, another copy of *which paper* is created, and it merges with the structure in (34) in order to check the strong *wh*-feature in C, forming (35).

\[(34)\quad [\text{CP did+Q [TP you [vP [vP [v} \text{v} [\text{vP file [which paper]i]] [\text{PP without PRO reading [which paper]i]]\\]
\]

\[(35)\quad \begin{array}{c}
\text{CP} \\
[\text{which paper}]^1-\text{WH} \\
\downarrow \\
C' \\
\downarrow \\
\text{did + Q} \\
\downarrow \\
\text{TP} \\
\downarrow \\
\text{you} \\
\downarrow \\
\text{T'} \\
\downarrow \\
\text{T} \\
\downarrow \\
\text{vP} \\
\downarrow \\
[\text{file [which paper]2-\text{WH} [\text{without PRO reading [which paper]3-\text{WH}]}}\]
\end{array}\]
At this point in the derivation, the copy of *which paper* in CP SPEC can form the chain \( CH_1 = (\text{copy}^1, \text{copy}^2) \), or the chain \( CH_2 = (\text{copy}^1, \text{copy}^3) \). In each case, the upper copy and the lower copy are nondistinct, and \( \text{copy}^1 \) c-commands \( \text{copy}^2 \) and \( \text{copy}^3 \). Furthermore, the wh-feature of \( \text{copy}^1 \) enters into a checking relation and the corresponding feature of the other copies is also accessible to the computational system, satisfying Last Resort. Finally, each chain satisfies the Minimal Link Condition, since there is no element containing a wh-feature that is closer to \( \text{copy}^1 \) than \( \text{copy}^2 \) and \( \text{copy}^3 \).

Therefore, Chain Reduction is applied to both chains \( CH_1 \) and \( CH_2 \), and the optimal reduction of \( CH_1 \) and \( CH_2 \) in the phonological component is obtained if the lower links of each chain are deleted, as shown in (36a,b). The structure (36b) is licit since each chain can be mapped into a linear order in accordance with the LCA. And thus, the derived sentence (37) is expected to be grammatical, as desired.

(36) a. \[
[\text{CP} \ [\text{which paper}]^{1-\text{WH}} \ \text{did}+\text{Q} \ [\text{TP} \ \text{you} \ [\text{vP} \ [\text{vP} \ \text{file} \ [\text{which paper}]^{2-\text{WH}}] \ [\text{PP} \ \text{without PROj reading} \ [\text{which paper}]^{3-\text{WH}]}]]]
\]

b. \[
[\text{CP} \ [\text{which paper}]^{1-\text{WH}} \ \text{did}+\text{Q} \ [\text{TP} \ \text{you} \ [\text{vP} \ [\text{vP} \ \text{file} \ [\text{which paper}]^{2-\text{WH}}] \ [\text{PP} \ \text{without PRO reading} \ [\text{which paper}]^{3-\text{WH}]}]]]
\]

(37) Which paper did you file without reading?

To summarize, the generalization that permits sideward movement is that given the nontrivial chains \( CH_1 \) and \( CH_2 \), the derivation is licit only if “(i) the links of \( CH_1 \) and the links of \( CH_2 \) are nondistinct; (ii) \( CH_1 \) and \( CH_2 \) have been formed before Spell-Out, and (iii) \( CH_1 \) and \( CH_2 \) have one link in common” (Nunes (2004: 105)).

What is interesting under the sideward movement analysis of parasitic gap constructions proposed by Nunes (1995, 2001, 2004) is that it can correctly account for various properties of parasitic gap constructions; however, in this section, I focus solely on the property called “S-structure Effects.”

It has been often noted in the literature (see Chomsky (1982)) that parasitic gaps are licensed at S-structure. The following contrast illustrates this point.

(38) a. Which paper did you file without reading?

b. *Who filed which paper without reading?

The ungrammaticality of (38b) indicates that covert movement of *which paper* cannot license the parasitic gap in the adjunct clause. Under the
framework of Minimalist Program, which postulates that the syntactic levels of representation are the interface levels LF and PF, this contrast is mysterious, and we must seek an alternative analysis. Nunes (2004) shows that the contrast above can be captured by the sideward movement analysis of parasitic gap constructions.

Let us now consider the derivation of (38b) in more detail. Under the relevant reading, *which paper* in the adjunct clause in (38b) ‘sideward moves’ from the object of *read* to the object of *file*, and the resulting structure is shipped to the phonological component by Spell-Out as shown in (39):

(39) [who [[filed [which paper]i] [without PRO reading [which paper]i]]]

Note that the two copies in (39) cannot form a chain because they do not satisfy the c-command requirement of Form Chain. If either copy of *which paper* were deleted, the linearization problem would not arise. However, Chain Reduction dictates that deletion for the purpose of linearization targets nontrivial chains, which means that the two copies in (39) are regarded as a trivial chain. And thus, Chain Reduction cannot be applied. Since neither copy of *which paper* was deleted in the phonological component, the chain violates the asymmetry and irreflexivity conditions on linear order, as desired. As we will see in section 4.2, I will reexamine the derivation of (38b) and point out a problem with this account.

To conclude, the sideward movement analysis and linearization of chains can successfully derive “S-structure Effects” of parasitic gap constructions without resorting to non-interface levels such as S-structure. This analysis derives not only “S-structure Effects,” but also the several properties of parasitic gap constructions such as the structural requirements that a parasitic gap cannot be c-commanded by the licensing gap (see Nunes (2004: Sec. 3.4)).

3.5. Summary

Nunes’s (2004) main proposal is that Move is not a primitive operation of the computational system, but rather it is a mere description of the interactions among the independent operations Copy, Merge, Form Chain, and Chain Reduction. If this is the case, the system should allow instances of sideward movement. Illicit instances of sideward movement are ruled out by the linearization of a nontrivial chain: if two copies of a given syntactic object do not form a chain, these copies are
regarded as an independent trivial chain, and Chain Reduction is inapplicable, which leads to a violation of the asymmetry and irreflexivity conditions on linear order. Finally, the sideward movement analysis derives the main properties of parasitic gap constructions.

4. Form Chain Reconsidered

One of the important consequences of the postulation of the Copy + Merge theory of movement is that movement of $\alpha$ takes place from a given syntactic tree to the other tree. In other words, the c-command requirement of movement should be dispensed with under the Copy + Merge theory of movement. As Hornstein (2001) explicitly notes, if movement is not constrained by c-command, as predicted by the Copy + Merge theory of movement, sideward movement is incompatible with Attract F proposed by Chomsky (1995). Therefore, under the view of the Copy + Merge theory of movement, movement is subject to Greed or Enlighten Self Interest (see Chomsky (1993), Lasnik (1995)).

Furthermore, under the theory developed by Nunes (2004), chains are representational syntactic objects, rather than just notational conventions (see also Takahashi (1994)). The derivational history of an element and its copy does not matter: a chain is legitimate if the links of the chain satisfies the c-command requirement at the end of the derivation. The links of a nontrivial chain must be deleted by Chain Reduction in order to linearize the syntactic objects. Chain Reduction requires that (at least) one link of the nontrivial chain must be phonetically realized, and it predicts that there is no case where no link of the chain is phonetically realized.

In this section, I will point out two problems with Nunes (2004). First, there are cases where a lower copy, not the head of the chain, is phonetically realized. Second, there are some instances where no member of a nontrivial chain is phonetically realized. In order to circumvent these problems, I will argue that chains are derivational syntactic objects as well as representational syntactic objects, which implies that c-command is not necessary under the definition of Form Chain in (25b).

4.1. Pronunciation of Lower Copies

Once we assume the copy theory of movement, we should expect that all members of chain links must be subject to the same principles.
Nunes (2004) convincingly argues that all chain links must satisfy the asymmetry and irreflexivity conditions on linear order in consonance with the LCA. One of the strong arguments for the copy theory of movement is from Bošković (2002), who argues that there is an exceptional instance where the lower copy of the wh-phrase is phonetically realized even in multiple wh-fronting languages, such as Serbo-Croatian. The relevant examples are shown below:

(40) a. *Šta šta uslovljava?
   what what conditions
   b. Šta uslovljava šta?
   what conditions what
   'what conditions what?'

(41) a. Šta neprestano šta uslovljava?
   what constantly what conditions
   b.*?Šta neprestano uslovljava šta?
   what constantly conditions what
   'what constantly conditions what?'

Bošković (2002), adopting Franks's (1998) analysis, argues that the lower copy of the wh-chain must be phonetically realized because of the phonological constraint that the two homophonous elements cannot be adjacent.

The question that arises is how we can determine the phonetic realization of a copy of the chain above under the framework of Nunes (2004). Note that the Chain Reduction, which deletes minimal number of copies of a chain in order to satisfy the linear order conditions, always chooses the head of a chain to be phonetically realized, since it participates in more checking relations than other copies. Let us consider the following structures of the sentences in (41) (English words are used for convenience):

(42)  [CP what\textsuperscript{1} [CP what\textsuperscript{2-WH} [VP constantly [VP what\textsuperscript{3-WH} [VP what\textsuperscript{4 [VP conditions what\textsuperscript{5-WH}]]]]]]]]

In (42), CH\textsubscript{1} = (what\textsuperscript{1}, what\textsuperscript{4}) and CH\textsubscript{2} = (what\textsuperscript{2}, what\textsuperscript{3}, what\textsuperscript{5}) are formed. Chain Reduction and FF-Elimination correctly determine that what\textsuperscript{1} in CH\textsubscript{1} is phonetically realized. Let us now consider CH\textsubscript{2}. In CH\textsubscript{2}, Chain Reduction and FF-Elimination predicts that what\textsuperscript{2} is the one that is phonetically realized. However, what\textsuperscript{2} cannot be phonetically realized due to the phonological constraint, and we would have to determine which copy of the chain CH\textsubscript{2} is phonetically realized. Unfortunately, each copy of what\textsuperscript{3} and what\textsuperscript{5} has the same features, and
thus Chain Reduction and FF-Elimination fails to choose the phonetic realization of the copy. The above discussion strongly suggests that Nunes (2004) would have to accept Franks’s (1998: Sec. 2.6.2) proposal that if the highest copy results in a PF crash, “economy considerations then dictates that the next highest copy is pronounced, unless again the result fails to converge.”

4.2. C-command Requirement

Even though the Copy + Merge theory of movement allows the targets of movement not to c-command the launching sites of that movement, the c-command requirement emerges under the condition of Form Chain: in a given chain CH = (α, β), α c-commands β. Under the definition of Form Chain in (25), the nontrivial chain cannot be created only in the case of sideward movement. In other words, sideward movement is special in that sideward movement does not create a chain. The question that arises is whether we can exclude the unwanted instances of sideward movement without resorting to the ‘special’ property of sideward movement.

In order to see why sideward movement does not create a nontrivial chain, let us reconsider the unacceptable structure in (39), which is repeated below for convenience.

(43) *Who filed which paper without reading?
(44) [who [[filed [which paper]%i] [without PRO reading [which paper]%i]]]

Under the structure, the wh-phrase sideward moves to the object position of the verb file in order to check the theta-feature, a kind of formal features. Nunes (2004: 108) argues that if sideward movement created

\[10\] As discussed by Nunes (2004), economy considerations regarding pronunciation of chain links are considerations such that pronunciation of a member of a nontrivial chain is more economical than pronunciation of two members of the chain.

\[11\] As correctly pointed out by an anonymous reviewer, under the analysis, the proposal by Franks (1998) is applicable only when Chain Reduction and FF-Elimination fail to choose pronunciation of the chain links. Thus, it seems that we need to assume a ranking regarding pronunciation of chain links. Nunes (2004: 37) also argues that “the ranking of best candidates for pronunciation from the highest to the lowest chain link follows from economy considerations regarding checking relations in overt syntax and application of FF-Elimination in the phonological component.”

\[12\] For helpful comments and suggestions, I wish to thank an anonymous reviewer.
a nontrivial chain, and if either copy of the wh-phrase were deleted, the resulting structure could be linearized and eventually converge. That is why he concludes that sideward movement does not create a nontrivial chain, and that the c-command requirement must be needed under the conditions of Form Chain.

Notice that under the mechanism of Nunes (2004), deletion of chain links is employed by Chain Reduction, which is subject to the FF-Elimination. Chain Reduction requires that all but one link of a chain must be deleted.

With this in mind, let us consider whether Chain Reduction can be applied to the structure in (44), assuming that in consonance with the standard view that chain represents the derivational history of movement, sideward movement also creates a nontrivial chain. The relevant tree diagram at some point of the derivation would be represented in (45):

\[
\begin{array}{c}
\text{TP} \\
\text{who} \\
\text{T'} \\
\text{T} \\
\text{vP} \\
\text{vP} \\
[\text{PP without reading [which paper]}'] \\
\text{filed} \\
\text{[which paper]}'
\end{array}
\]

Note that in (45), there is no element that asymmetrically c-commands the copy of which paper, and in turn is asymmetrically c-commanded by the copy. This clearly indicates that each copy in (45) satisfies the asymmetry and irreflexivity conditions on linear order. In other words, each copy in (45) will be treated by the phonological component as a trivial chain and thus, Chain Reduction must not be inapplicable to the structure. Hence, even if the copies in (45) form a (nontrivial) chain, the deletion of either copy in (45) cannot be applicable to the structure. Once one of the copies was deleted, the derivation would be excluded, as shown in (43).

To put it another way, the reason why Chain Reduction cannot apply to the structure in (45) is not due to the fact that the member of the chain does not c-command the other, which is a requirement for Form Chain, but due to the fact that the structure does not violate the asym-
metry and irreflexivity conditions on linear order. Therefore, although the c-command requirement under the definition of Form Chain is dubious, we can still keep the original conception of chains, namely, a chain represents a derivational history of movement.\(^{13}\)

If sideward movement forms a (nontrivial) chain, as discussed above, how can we account for the unacceptable sentence in (46)? The relevant structure of (46) would be given in (47):

\[(46) \quad \ast \text{Who filed which paper without reading?}\]

\[(47) \quad [\text{who} \begin{array}{c} \text{[filed} [\text{which paper}]\text{]} \\ \text{[without PRO reading} [\text{which paper}]\text{]} \end{array}]\]

Recall that the ungrammaticality of (46) is not due to the fact that each member of the chain fails to c-command the other. Remember that Chain Reduction dictates that all but one link of a nontrivial chain must be deleted unless morphological fusion makes the copy invisible for the LCA. If this is the case, then the structure in (47) is correctly excluded without resorting to the c-command requirement of Form Chain. As discussed above, the chain in (47) does not violate the asymmetry and irreflexivity conditions on linear order, and thus, Chain Reduction cannot be applied to the structure. However, the derivation of (46) is excluded because of the fact that one of the chains is incorrectly deleted by Chain Reduction. If this is the case, no deletion of each copy in (47) should be acceptable, which is borne out by the following example:

\[(48) \quad \text{Who filed which paper without reading which paper?}\]

Since the instances of \textit{which paper} in (48) do not form a nontrivial chain, they are regarded as a trivial chain. Thus, Chain Reduction is inapplicable and they must be phonetically realized.

To summarize so far, I have shown that even though sideward movement creates a nontrivial chain, contrary to Nunes (2004), the unwanted derivations can still be excluded. Therefore, we may maintain the original view that chain represents a derivational history of movement.\(^{14}\)

\(^{13}\) This argument still applies even though we assume the derivation in which the \textit{wh}-phrase in (45) overtly moves to vP SPEC, where the moved copy c-commands the copy in the adjunct clause. However, as Nunes (2004: 108–109) explicitly argues, the moved copy and the copy in the adjunct cannot form a chain because of Last Resort and Minimal Link Condition.

\(^{14}\) There is another possibility to account for parasitic gaps in English without
4.3. More on C-command Requirement

So far, I have shown that illicit instances of sideward movement can be captured even if we assume that sideward movement forms a nontrivial chain. In this subsection, I will provide evidence that sideward movement must form a nontrivial chain.

As Nunes (2004: 27) explicitly notes, “recoverability of deletion in the case of Chain Reduction can be ensured by economy considerations regarding the length of derivations.” Thus, the generalization that Chain Reduction makes is that (at least) one member of a nontrivial chain must be phonetically realized (see also Hornstein (2001: 80)).

Note, however, that it seems that all members of the nontrivial chains in (49) are fully deleted. The example is from Bošković (2002), who presents compelling evidence for the copy theory of movement.

(49) Romanian (from Bošković (2002))

Ce Precede ce fara sa influenteze?
what precedes what without SUBJ.PRT influence-3P.SG
‘What precedes what without influencing?’

What is interesting here is that the wh-in situ in (49) can license the parasitic gap. Under the standard analysis that the wh-in situ does not license parasitic gaps, the sentence in (49) is totally unexpected. Bošković (2002) argues that since Romanian is one of the languages that allows multiple wh-fronting, the wh-in situ actually moves to a higher position, where the parasitic gap in the adjunct clause is licensed. However, due to the phonological constraint that two homophonous element must not be adjacent, the wh-phrase ce in (49) must be pronounced in the base position.

Let us now consider the structure in (49) in more detail and see how the sideward movement analysis accounts for the parasitic gap in (49) (English words are used for convenience).

(50) $[[\text{CP what}^1 \ [\text{CP what}^2 \ [\text{vP what}^3 \ [\text{vP precedes what}^4 \ [\text{PP without influencing what}^5]]]]]]$

\[\text{resorting to sideward movement. Nissembaum (2000) argues that parasitic gaps in English can be accounted for by covert movement. However, to discuss covert movement analysis of parasitic gaps in English as a whole is beyond the scope of this paper. See Nissembaum (2000) for further details of parasitic gaps in English.}\]

\[\text{\footnote{The exception here is only when the link of a nontrivial chain undergoes morphological fusion. See section 2.6.}}\]
The \textit{wh}-phrase \textit{what}$^5$ in the adjunct clause sideward moves to the object position of the verb \textit{precedes}.\textsuperscript{16} The subject \textit{wh}-phrase \textit{what}$^3$ also undergoes movement to CP SPEC and finally the \textit{wh}-phrase \textit{what}$^4$ undergoes movement to CP SPEC in overt syntax. Under the analysis of Nunes (2004), the structure in (50) has $\text{CH}_1 = (\textit{what}^1, \textit{what}^3)$, $\text{CH}_2 = (\textit{what}^2, \textit{what}^4)$ and $\text{CH}_3 = (\textit{what}^2, \textit{what}^5)$. Chain Reduction applies to each chain, and \textit{what}^3 in $\text{CH}_1$ is deleted in order to satisfy the asymmetry and irreflexivity conditions on linear order. Chain Reduction also applies to $\text{CH}_2$, but the phonological constraint dictates that \textit{what}^2 cannot be phonetically realized. Thus, \textit{what}^4 must be phonetically realized, as desired. Finally in the case of $\text{CH}_3$, \textit{what}^5 would be deleted under Chain Reduction, which is the same as parasitic gap constructions in English (see section 3.4). Note that in Romanian, \textit{what}^2 cannot be phonetically realized because of the phonological constraint. Thus, the theory under Nunes (2004) would predict that \textit{what}^5 instead of \textit{what}^2 must be phonetically realized. However, the grammatical example in (49) clearly indicates that \textit{what}^5 is also deleted, which means that all members of the chain $\text{CH}_3$ are deleted. This contradicts the generalization predicted by Chain Reduction; namely (at least) one member of a nontrivial chain must be phonetically realized.

4.4. Toward an Account

In this subsection, I will discuss one possibility to account for the sentence in (49). Recall that Chain Reduction requires that (at least) one link of a chain must be phonetically realized. The problem with the Romanian example in (49) is the fact that no member of $\text{CH}_3$ is phonetically realized, which is a violation of the requirement of Chain Reduction. This problem is avoided only if sideward movement creates a nontrivial chain. With this in mind, consider the structure in (50), which is repeated below:

\begin{equation}
\text{(51) } [\text{CP } \textit{what}^1 \text{ [CP } \textit{what}^2\text{-WH [vP } \textit{what}^3 \text{ [vP precedes } \textit{what}^4\text{-WH]} \text{ [PP without influencing } \textit{what}^5\text{-WH]}]]]
\end{equation}

Under the view, $\text{CH}_2 = (\textit{what}^2, \textit{what}^4, \textit{what}^5)$ is formed, in addition to $\text{CH}_1 = (\textit{what}^1, \textit{what}^3)$. The question that arises is which link of

\textsuperscript{16} I will simply follow Nunes (2004), who argues that sideward movement always takes place from a ‘subordinated’ clause to a ‘subordinating’ clause. See Section 3.7 of Nunes (2004) for detailed discussion on that matter.
the chain CH$_2$ is phonetically realized. Obviously, what$_2$ cannot be phonetically realized because of the phonological constraint, following Bošković (2002). Thus, we have to choose which copy of the chain is phonetically realized. It seems that Franks's (1998) proposal concerning pronunciation of lower copies does not work. Franks (1998) proposes that if the pronunciation of the head of the chain causes a PF crash, the next highest copy is pronounced. It is not clear whether or not what$_4$ is the higher copy than what$_5$. Let us suppose that economy considerations dictate that the copy that is the most related to the highest one is pronounced.\textsuperscript{17} With this in mind, let us consider the following structure of the sentence in (49):

\[(52) \quad [\text{CP what}^1 [\text{CP what}^2_{-WH} [\text{VP what}^3 \text{VP precedes what}^4_{-WH}] [\text{PP \Theta}1, \Theta2 \text{\ without influencing what}^5_{-WH}])]\]

\[\Theta1\]

Let us assume that sideward movement takes place from a subordinated domain to a subordinating domain (see Nunes (2004)). Under the derivation, what$_5$, which checks the theta feature of the verb influencing, leaves a copy, and the copy merges with the verb precedes to check its theta feature. Thus, what$_4$ has two theta features. After that, the copy of what$_4$ undergoes movement to CP SPEC. Notice that the copy of what$_5$ cannot move to the matrix CP SPEC, since it crosses an island (See Nunes (2004: Sec. 3.7), and Nunes and Uriagereka (2000)). Since what$_2$ and what$_4$ have two theta features, while what$_5$ has only one, we may say that what$_4$ is the most related copy to what$_2$, and thus what$_4$ would be phonetically realized, as desired.

To conclude this subsection, I have discussed one possibility to account for the sentence in (49). Needless to say, that is certainly not the only possibility, and the exact mechanism of the phonetic realization of lower copies must be worked out in more detail. However, it is now clear that chain is derivational and representational syntactic objects, and that additional mechanisms must be added to the framework of Nunes (2004).

\textsuperscript{17} This account needs to assume another ranking regarding pronunciation of chain links. I am indebted to an anonymous reviewer for pointing this out to me.
4.5. Summary

In this section, I suggested that chain formation characterizes a derivational history of movement as well as representational syntactic objects. As a consequence, the c-command requirement is not necessary for the definition of Form Chain in (25). Under this view, chain is created every time when an element undergoes movement. In this sense, sideward movement is not a 'special' operation, as it can also create a nontrivial chain. Note that I assume the Copy + Merge theory of movement: namely, Move is not a primitive singulary operation of the computational system. Thus, even though movement and chain formation express the same type of relation, Brody's (1995) criticism can be circumvented.

5. Concluding Remarks

In the book under review, Nunes (2004) attempts to formalize the displacement property of human language following copy theory of movement proposed by Chomsky (1995). Nunes proposes that movement of syntactic objects is carried out by four independent operations: Copy, Merge, Form Chain, and Chain Reduction. Chains created by movement are subject to the LCA, and thus the copies of the chain must be deleted in accordance with the LCA. In this way, the theory correctly answers the questions of why (in general) all members of the links cannot be phonetically realized, and why (in general) the tail of the chain is not phonetically realized. Furthermore, he argues that sideward movement, derived from the Copy + Merge theory of movement, is empirically motivated. In this review article, I have made a modification concerning Form Chain: chain formation does not require c-command. If c-command is not needed, we can maintain the standard view of chains, which states that chains represent a derivational history of movement. I have also shown that unwanted instances of sideward movement are also excluded by Chain Reduction.

REFERENCES


Bobaljik, Jonathan and Samuel Brown (1997) "Inter Arboreal Operations: Head

Department of Early Childhood Education and Care
Kochi Gakuen College
292–26 Asahitenjincho, Kochi,
Kochi 780–0955
e-mail: niinuma@kochi-gc.ac.jp