

HOW STRONG ISLANDS ARE DERIVED FROM THE WAY A TOP-DOWN DERIVATION IS LINEARIZED

Valérie Gautier

University of Nantes

The goal of this paper is to show that strong islands can be derived from the way derivation is linearized, as long as we assume that the derivation proceeds in a top-down fashion.

To begin with, I will present one advantage of adopting a top-down approach regarding linearization issue: the Linear Correspondence Axiom (henceforth, LCA, Kayne 1994) can be reformulated in a more derivational and minimalist way. In particular, Kayne assumes that the notion that derives precedence is the asymmetric c-command. Because of the asymmetric c-command, the LCA rules out the head-complement configuration in bare phrase structures. I show that with a top-down derivation, the problematic asymmetric c-command relation can be eliminated. With our *Top-down LCA*, precedence relations are derived from the way phrase structures have been built: roughly, if Y enters the derivation right after X, then X precedes Y. In a second part, I will argue that this non-standard approach to derivation¹ and linearization can capture CED effects: it will be demonstrated that subjects and adjuncts are islands because they have to be built in a parallel derivation.

1. A top-down derivation

In the generative framework, it is taken for granted that derivations proceed bottom-up. Roughly, a bottom-up derivation starts with the structurally lowest constituent in a clause and ends with the structurally highest one (such as subject or fronted elements). As an illustration, consider the derivation of a simple sentence like (1).

- (1) La fille danse.
The girl dances.

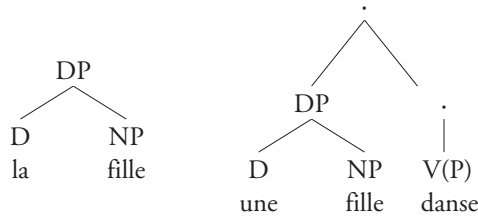
¹ The idea that a top-down derivation can derive CED effects has been first proposed by Boeckx (1999).

(2)

V(P) is the first
element entering
the derivation
V(P) danse

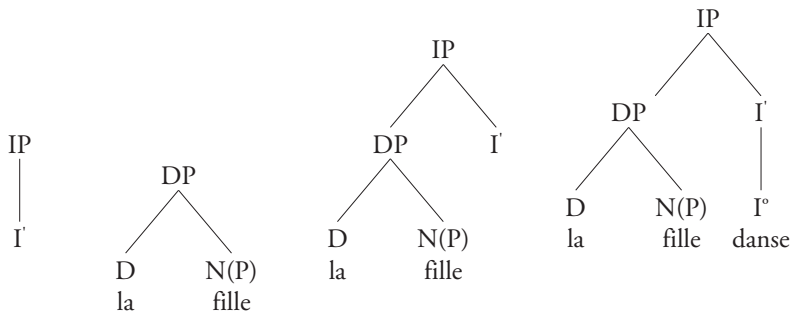
The subject DP is
assembled in parallel.

DP is merged with VP



Following Phillips (1996, 2001), I claim that the derivation should proceed in a top-down fashion. This non-standard assumption entails that the derivation of (1) starts with I(P) and ends with V(P). The top-down derivation of (1) is given below.

- a) IP-I' is the first element entering the derivation
 b) To check I' nominative features, DP subject is needed. [la fille] is then built in parallel
 c) DP is merged with I'. Nominative features are checked
 d) The inflexion enters the derivation.



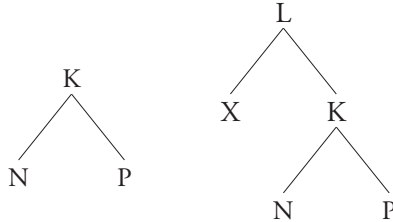
Bottom-up vs Top-down Merge

From a minimalist perspective, the phrase structure of a given sentence is built by the two structure-building operations *Merge* and *Move*. In a bottom-up approach, *Merge* is assumed to apply at the root² of the tree, that is, *Merge* is a cyclic operation. (3) is the definition of cyclic *Merge*.

² « Merge always applies in the simplest possible form: at the root ». (Chomsky 1995)

(3) *Cyclic Merge*

Merge X and K yields the new constituent L. L dominates the tree.

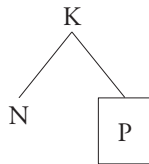


In a top-down derivation, *Merge* crucially does not increment the tree at the root: each time a new item is inserted in the derivation, it replaces something in the structure. *Merge* in a top-down derivation is then a counter-cyclic operation. I define this counter-cyclic operation as *Top-down Merge*. *Top-down Merge* is given in (4) and illustrated in (5) and (6).

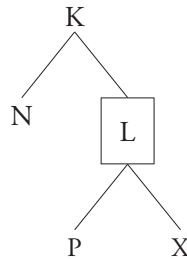
(4) *Top-down Merge*

- a) Applied to α and β , Merge forms the new object $\{L, \{\alpha, \beta\}\}$ by merging α and β .³
- b) L replaces the last terminal node merged in the structure.

(5)



(6)



As an illustration, let us see how the VP in (7) is concatenated.

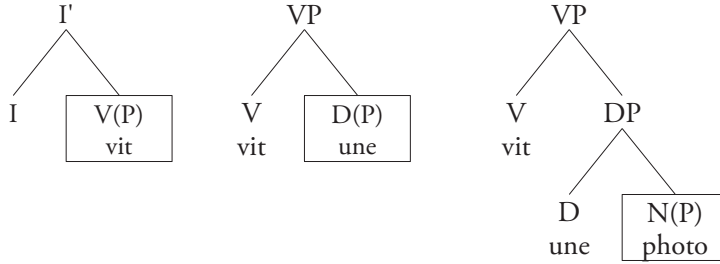
(7) [vit une photo de Marie]
saw a picture of Mary

1
At this step, V(P) is the last terminal node merged

2
D(P)une enters the derivation
a) VP [vit une] is built,
b) VP [vit une] replaces V(P) [vit]

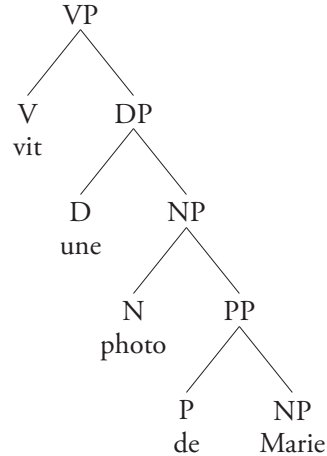
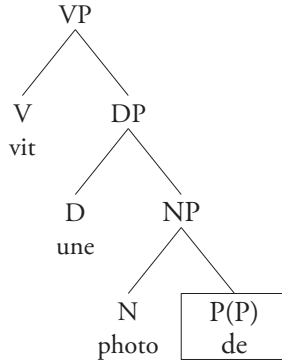
3
N(P)photo enters the derivation
a) DP [une photo] is built
b) DP [une photo] replaces D(P)une

³ This definition has been inspired by the definition of counter-cyclic *Merge* given by Kitahara (1995).



- 4
 P(P)*de* enters the derivation
 a) NP [photo de] is built,
 b) NP [photo de] replaces N(P) [photo]

- 5
 N(P)*Marie* enters the derivation
 a) PP [de Marie] is built,
 b) PP[de Marie] replaces P(P)*de*



Several arguments have already been given in favor of a top-down syntax.⁴ In particular, Phillips (1996, 2001) showed how a top-down derivation⁵ explains why constituency tests sometimes yield contradictory results. As Phillips pointed out, building the sentence in a top-down way entails that the insertion of an item in the derivation can destroy the preceding constituent to create a new one.⁶ That is, contrary to a bottom-up approach, constituent structure is not permanent. As a consequence, when the constituency test's contradictions are mysterious in a bottom-up approach, there are expected in a top-down derivation.

As an illustration, consider (8) and (9) below.

⁴ See Boeckx (1999), Richards (1999), Guimaeres (2004), Drury (2005), for instance.

⁵ Note that Phillips also assumes that the derivation proceeds from left-to-right. I am not making this assumption here

⁶ For instance, in the derivation of (7), the insertion of the N(P)*Marie* destroys the constituent NP*photo de*.

(8) John *talked to* and *gossiped about* the kid who sprayed paint on his car.

(9) Helen *talked to* Jonathan and Alice did * (to) Matthew.

Building the structure in a top-down way entails that in (8), [talked to] is still a constituent as it has not been merged with a argument. Thus, [talked to] can be coordinated. In contrast, in (9), the addition of *Jonathan* causes destruction of the constituent [talked to]. As the verb and the preposition can no longer form a unit, [talked to] cannot be deleted in (9) (for details, see Phillips 2003). In the next section, I'm going to present another strong argument for a top-down syntax. One advantage of a top-down derivation is that it provides us a way to reformulate the LCA into an axiom which fits in a minimalist approach.

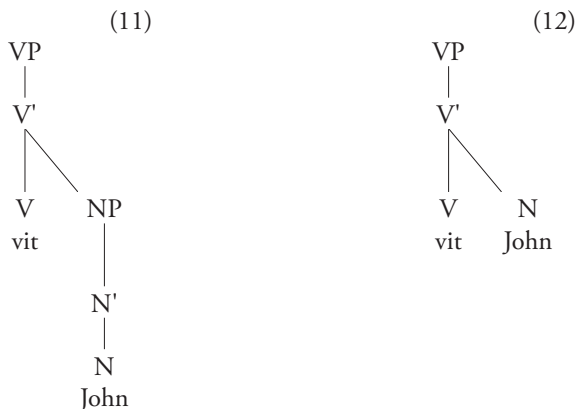
2. Why the LCA should be reformulated

Kayne (1994) argues that there should exist a correspondence between hierarchical relations in a phrase structure and the linear order between terminal nodes. Kayne's idea is that this correspondence is given by the asymmetric *c*-command relation. He formalized this intuition with the LCA.

(10) Linear Correspondence Axiom (LCA)

A lexical item α precedes a lexical item β if α asymmetrically *c*-commands β .

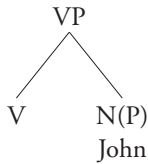
It should be reminded to the reader that Kayne's proposal has been elaborated in the Government and Binding framework. Kayne argues that (10) derives the properties of one of the central GB's module: X-bar theory. Crucially, the LCA justifies the fact that in X-bar schemata, the complement of a head cannot be another head. As an illustration, let us compare (11) and (12). In (11), the complement of the verb is the maximal projection NP, with the three X-bar levels. Note that V *c*-commands N, but N does not *c*-command V. By (10), V precedes N. (11) is a licit configuration for the LCA. In (12), the complement of the verb is the head N° *John*. Here, no order can be established between V and N because V and N are not in an asymmetric relation (V and N being dominated by the same node). (12) is illicit for the LCA.



Since GB model, progress in the field of constituency description has been made, leading to a discussion of the adequacy of the X-bar template. In particular, the relevance of intermediate vacuous projections, such as N' in (11), is questioned.

As a consequence, in the minimalist framework, X-bar theory is replaced by the bare phrase structure. In this new approach to phrase structure, lexical items selected in the numeration, have a double status: they are minimal and maximal projections. This means that the head-complement configuration in (11) is replaced by the head-complement configuration in (13), where *John* is dominated by a minimal and maximal projection node.

(13)



Note that in adopting (13), the head-complement configuration becomes a phrase structure ruled out by the LCA. That is, (13) is (12): the complement of the head V is another head.

In brief, in the X-bar theory, a head's complement has to be a maximal projection. With the asymmetric c -command relation, the LCA derives this property. In the bare phrase structure, a head's complement can be another head. By ruling out this syntactic configuration, the LCA does not derive this property. I conclude that the LCA is not an appropriate principle for the new phrase structure formalism adopted by the generative model. To solve this compatibility problem, I suggest the reformulation of the LCA. The reasoning is the following: as X-bar theory has been replaced by bare phrase structure, the LCA should be replaced by a new Linear Correspondence Axiom, which fits with the bare phrase structure formalism.

3. Top-down LCA

To see what should be changed in the LCA, we need to know what in Kayne's axiom, repeated in (14), makes the structure in (13) illicit.

(14) Linear Correspondence Axiom (LCA)

A lexical item α precedes a lexical item β if α asymmetrically c -commands β .

(13) is bad because V and N stand in a mutual c -command relation. It just so happened that the relation which derives precedence is the asymmetric c -command relation. What needs to be changed seems quite obvious: the asymmetric c -command should be eliminated from the LCA. What should be then the relation that maps hierarchical relations onto linear order? Note that the simple c -command relation is not good either since V c -commands N and N c -commands V . I suggest that in order to reformulate the LCA into a satisfying minimalist principle, we need a deriva-

tional axiom which says that V precedes N because V entered the derivation before N. The only way to get such an axiom is to assume that the derivation proceeds in a top-down way. I propose (15) as the reformulation of the LCA.

(15) *Top-down LCA*

At the derivational step n , a lexical item α precedes immediately a lexical item β iff

α c -commands β

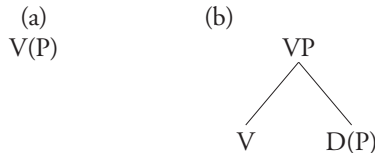
β is the last terminal node merged at step n

α was the last terminal node merged at step $n - 1$.

NB: β and α are terminal nodes

(15) entails that c -command relations map onto precedence relations in a derivational way: at each derivational step, the last terminal node α merged in the tree precedes the new terminal node β inserted in the derivation. With (15), we do not need the asymmetric c -command anymore: when a term β arrives in the derivation after a term α , then β cannot precede but will automatically follow α (even if, representationally, α and β are in a mutual c -command).

It is shown in a) and b) below how (15) derives the precedence relation for a head-complement configuration: The complement $D(P)$ follows the head V , because the lexical item $D(P)$ has been inserted in the derivation after the lexical item $V(P)$.



As a concrete illustration, let's go back to the derivation of [vit une photo de Marie], to see how the new LCA works.

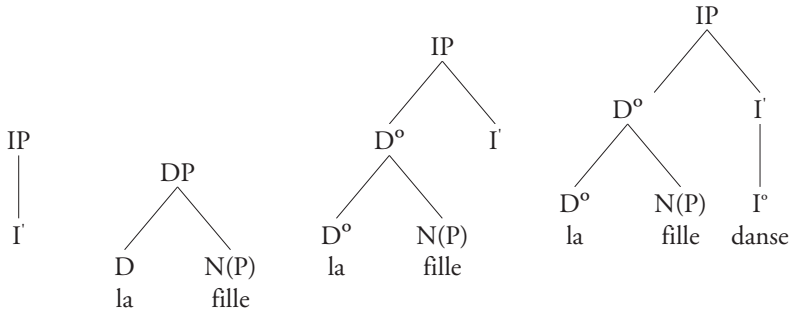
Step 2: the terminal node $D(P)une$ is merged with the terminal node $V(P)vit \rightarrow D(P)une$ follows $V(P)vit$. Step 3: the terminal node $N(P)photo$ is merged with the terminal node $D(P)une \rightarrow N(P)photo$ follows $D(P)une$. Step 4: the terminal node $P(P)de$ is merged after the terminal node $N(P)photo \rightarrow P(P)de$ follows $N(P)photo$. Step 5: the terminal node $N(P)Marie$ is merged after the terminal node $Pde \rightarrow N(P)Marie$ follows Pde . At this point of the discussion, we have demonstrated how (15) derives linear order for terminal nodes in a head-complement configuration. I left aside the question of linearization of specifiers and adjuncts. In the second part of this paper, this problem will be considered. In particular, I am going to show that specifiers and adjuncts have a special status regarding linearization, that is, when they enter the derivation, (15) cannot apply. I will argue that the Constraint on Extraction Domain (CED, Huang 1982) can be derived from this failure.

4. Why adjuncts and subjects are islands: a top-down explanation

The specifier position will be the first case examined here. Let us go back to the top-down derivation of (1), repeated in (16).

- (16) Une fille danse.
The girl dances.

- a) IP-I' is the first element entering the derivation b) To check I' nominative features, DP subject is needed. [la fille] is built in a parallel derivation. c) DP is merged with I'. Nominative features are checked d) The inflexion enters the derivation.



Note that before entering the derivation, DP has been spelled-out to become a D°.⁷ The “reduction” of the maximal projection into a terminal node is necessary given the assumption that only terminal nodes can enter a top-down derivation.⁸

At step c), the ordering relation between DP and I' is unspecified. Remember that (15) gives the precedence relation between two terminal nodes: since I' is not a terminal node, (15) cannot apply. The spec-head order is finally obtained at step d), once I° enters the derivation: D° c-commands I° and I° has entered the derivation after D°, then D° precedes I°. I claim that the way the subject is built in a top-down derivation explains subject's islandhood. It is well-known, that DPs cannot be extracted from a subject position. As an illustration, consider (17).

- (17) *Who_i does a picture of t_i upset Mary ?

Remember that the subject is built in a parallel derivation. This entails that the wh-phrase “who” and the subject DP “a picture of” belong to two distinct derivational workspaces. “who” belongs to (i) and “a picture of” belongs to (ii).

⁷ This assumption is in the spirit of Uriagereka and Nunes (1999)'s Multiple Spell-out Model.

⁸ The introduction of IP-I' being an exception in this respect.

- (i) */Who does /*SUBJECT*/ upset Mary/
 (ii) /a picture of/

I suggest that DPs must check their thematic features in their own derivational workspace.⁹ This proposal is formulated in (18).

- (18) A DP must check its theta-role (see Hornstein 1998) in its workspace. Otherwise, the derivation crashes.

In (17), —“who” cannot check its theta-role with “upset”, since “upset” already has an argument. The only possibility for “who” to check its theta-role is then with the preposition “of”. As “who” and “of” do not belong to the same workspace, the derivation crashes. The ungrammaticality of (17) is correctly predicted by (18). Note that (18) predicts that (17) should be good if “upset” has no overt argument. This prediction is correct: (19) is grammatical because the complement position of “upset” has been filled with a trace. The reason why a parasitic gap construction legitimates an extraction out of a subject is then intuitively explained.

- (19) Who_i does a picture of t_i upset t_i ?

As there is no overt argument for “upset”, a copy of “who” can be inserted in complement’s position of “upset”. “who” can then check its theta-role in its derivational workspace, the derivation does not crash. A copy of “who” is created in the complement position of “of” to check the selectional properties of the preposition.

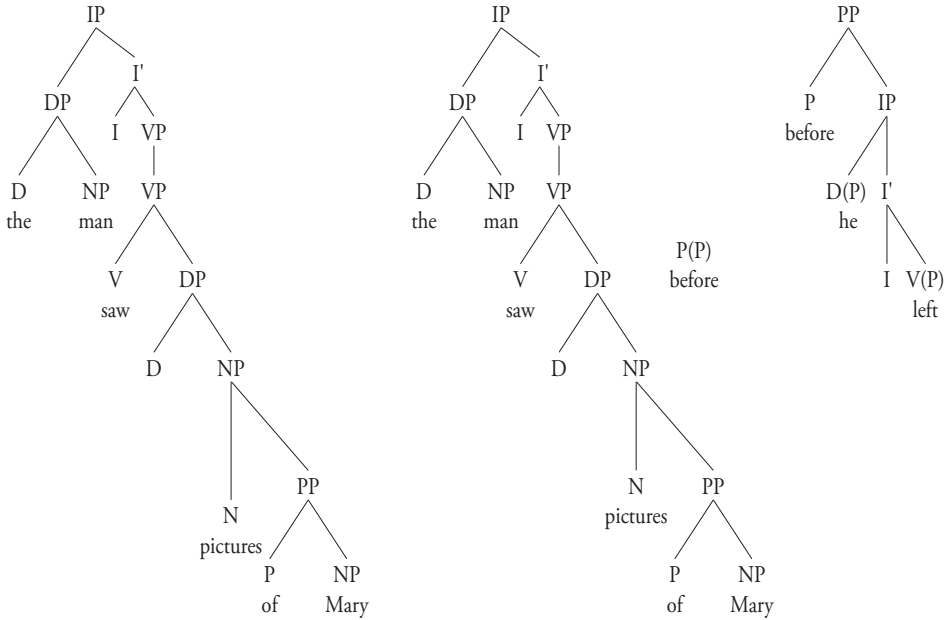
- (i) /Who does /*SUBJECT*/ upset *whol*
 (ii) /a picture of *whol*

We have just seen how the way subjects are introduced in a top-down derivation explains why they are strong islands. Let us examine now how (15) deals with adjuncts. For the sake of the presentation, we will consider only cases relevant to the island paradigm: that is, the discussion will be limited to right adjunctions. Note that, contrary to Kayne’s proposal, right adjunction is a legitimate operation here. (20) is an instance of right adjunction.

- (20) The man saw pictures of Mary before he left.

- | | | |
|--|--|--|
| a) VP has been built. The last terminal node merged in the structure is N(P) <i>Mary</i> | b) P(P) enters the derivation. P(P) cannot replace N(P). | c) In order to prevent of <i>Top-down LCA</i> ’s violation: → PP is built in parallel. |
|--|--|--|

⁹ Crucially note that here, sideward movement (along Nunes’ 1995 lines) is not allowed.



At step b) “before” is the new element to be inserted in the derivation, after [the man saw pictures of Mary] has been built. The terminal node $P(P)_{\text{before}}$ cannot be merged with the last lexical head arrived in the derivation (“Mary”). As a consequence, no ordering can be established between between $P(P)$ and $D(P)$ “Mary” (since $D(P)$ does not c-command $N(P)$). As a result, the adjunct is built in parallel.¹⁰ The fact that the adjunct is built in a derivational workspace distinct from the main derivation explains sentence like (21). (21) is bad because a *wh*-phrase has been extracted from the adjunct [before filing].

- (21) *Which paper_i did you read Don Quixote before filing t_i ?
 (i) */Which paper did you read Don Quixote/
 (ii) /before filing/

[which paper] belongs to the derivational workspace (i).¹¹ Since “read” already has an argument, [which paper] cannot check its theta features in its own derivational workspace. As a consequence, (18) is violated and the derivation crashes. Adjunct’s islandhood is then justified. As for subject cases, (21) becomes good with a parasitic gap construction.

¹⁰ The question of how the adjunct is finally linearized will not be discussed in this paper.

¹¹ Even if the DP, as a spec-CP, has been built in a parallel derivation. As the end, it belongs to the “main” derivation.

- (22) Which paper did you read t_i without filing PG_i ?
 (i) /Which paper did you read *which paper*/

As there is no overt external argument for “read”, “which paper” can check its thematic features: the derivation does not crash. A copy of “which paper” is created in (ii) and inserted to check selectional properties of “filing”.

- (ii) /before filing *which paper*/

Conclusion

I showed that adopting a top-down derivation allows us to reformulate the LCA in a derivational way, and to eliminate the asymmetric c-command relation. The way our Top-down LCA applies provides us with a story to derive straightforwardly subjects and adjuncts’ islandhood: nothing can be extracted from these constituents because they are built in a parallel derivation. The reason why movement across islands becomes acceptable when a “copy” is inserted in the parallel derivation (i.e. parasitic gap) is then intuitively justified.

References

- Boeckx, C., 1999, *Programmatic Notes for a Top-down Syntax*.
 Chomsky, N., 1995a, *Bare phrase structure*. In *Government and binding theory and the Minimalist Program*, ed. G. Webelhuth, 343-439. Oxford: Blackwell. [Reprinted in *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, ed H. Campos & P. Kempchinsky, 51-109. Washington D.C.: Georgetown University Press].
 — 1995b, *The Minimalist Program*. MIT Press, Cambridge Massachusetts. Linguistics.
 — 2000, «Minimalist Inquiries: The framework». In *Step by step: Essays in minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka, 89-155. Cambridge, Mass.: MIT Press.
 Drury, J., 2005, *Alternative Directions for Minimalist Inquiry: Expanding and Contracting phases of the derivation*. Doctoral Dissertation. University of Maryland.
 Fox, D. & Pesetsky D., 2004, «Cyclic linearization of syntactic structure». To appear in *Theoretical Linguistics, special issue on Object Shift in Scandinavian*; Katalin E Kiss, ed.
 Gautier, V., 2000, *Dérivation linéaire et Cyclicité*. Master thesis. University of Nantes.
 Guimaraes, M., 2004, *Derivation and representation of syntactic amalgams*. Doctoral dissertation. University of Maryland.
 Huang, C.T. J., 1982, *Logical relations in Chinese and the theory of grammar*.
 Kayne, R., 1994, *The Antisymmetry of syntax*. MIT Press. Cambridge Massassuchets.
 Kitahara, H., 1995, «Target α : deducing Strict Cyclicity from Derivational Economy», *LI*.
 Murgja, E., 2004, *Syntactic Identity and Locality Restrictions on Verbal Ellipsis*. Doctoral Dissertation. University of Maryland.
 Nunes, J., 1995, *The copy theory of movement and linearization of chains in the minimalist program*. Doctoral Thesis. University of Maryland.
 Phillips, C., 1996, *Order and Structure*. Doctoral Dissertation. MIT.
 — 2003, «Linear Order and constituency», *LI* 34, 37-90.

- Richards, N., 1999, «Dependency Formation and Directionality of Tree Construction». In *MITWPL 34: Papers on Morphology and Syntax*, Cycle two, Vivian Lin, Cornelia Krause, Benjamin Bruening and Karlos Arregi, eds. Cambridge: MIT Working Papers in Linguistics.
- Stepanov, A., 2001, «Late adjunction and minimalist phrase structure», *Syntax* 4, 94-125.
— Doctoral Dissertation. MIT, Cambridge, Massachusetts.
- Uriagereka, J. & Nunes, J., 2001, «Cyclicity and Extraction domains», *Syntax* 3, Number 1. April 2000.