

## Chain Formation as inclusion relations between copies

Carlos Muñoz Pérez

Universidad de Buenos Aires & Georg-August-Universität Göttingen

cmunozperez@filo.uba.ar

### 1. Three explanatory challenges introduced by Copy Theory

Take the sentences in (1).

- (1) a. John was kissed.  
b. John kissed John.

And now consider their structure according to the *Copy Theory of Movement* (Chomsky 1993). Remember: under Copy Theory movement does not exist as an operation (i.e., nothing really moves), but it is an epiphenomenon of how the PF component phonetically realizes *chains*<sup>1</sup> consisting of *non-distinct* (i.e., “identical”) copies.

- (2) a. [TP John [T<sup>1</sup> was [VP kissed John]]]  
b. [TP John [T<sup>1</sup> T [VP John [V<sup>1</sup> kissed John]]].

Thus, Copy Theory should provide us with a principled explanation on how we get the “phonological” representations in (1) from the syntactic structures in (2). This explanation must be capable of dealing with three different problems posed by Copy Theory. Let’s call them the *Challenges of Copy Theory*.

- (3) *Challenges of Copy Theory*  
a. *Sameness Challenge*  
How can we tell when two occurrences of the same lexical item are copies and when they are not?  
b. *Uniqueness Challenge*  
Why is it necessary to delete all copies but one?  
c. *Rank Challenge*  
Why is it forbidden to delete the highest copy?

<sup>1</sup> Adapting somewhat Chomsky’s (1981) terminology, a chain CH is a sequence (Copy<sup>1</sup>, ..., Copy<sup>n</sup>) which represents the derivational history of a constituent by successive applications of the operations *Copy* and *Merge*.

In what remains of the presentation, I will show what I think is wrong with the “standard” approach to these questions (Nunes 1995, 2004). Based on the problems of the standard account, I am going to propose a novel approach to the three challenges. Finally, I am going to show how this new system can deal with some of the exceptions to the Uniqueness and Rank Challenges.

### 2. Nunes’ (1995, 2004) answers

The first fine grained and principled theory trying to answer the explanatory challenges of Copy Theory was offered by Nunes (1995, 2004). However, at least two of the answers posed by Nunes suffer important drawbacks.

#### 2.1. THE SAMENESS CHALLENGE

The most important property of a movement chain under Copy Theory is *Non-Distinctiveness*: all the links in the chain should be the “same” in some relevant way. However, defining Non-Distinctiveness is not that easy.

- (4) [TP John<sub>1</sub> [T<sup>1</sup> T [VP John<sub>2</sub> [V<sup>1</sup> said that [TP John<sub>3</sub> [T<sup>1</sup> was [VP kissed John<sub>4</sub>]]]]]]].

What kind of principle allows identifying as “the same” *John*<sub>1</sub> and *John*<sub>2</sub> and as distinct *John*<sub>2</sub> and *John*<sub>3</sub>? In other words, what kinds of elements are computed as *identical*?

Given the difficulty of distinguishing real syntactic copies from occurrences of the same lexical item, Chomsky (1995) proposed a marking mechanism to signal those elements related by internal Merge (just as in *Trace Theory*). Nunes follows this idea:

I will follow Chomsky’s (1995, 227) proposal that two lexical items *l* and *l*’ selected from a numeration *should be marked as distinct* for the computational system if they are accessed by distinct applications of Select. As for phrasal objects, I assume that their labels encode the relevant piece of information regarding distinctiveness; the DP [the man], for instance, is to be represented in bare phrase structure terms as  $K = \{\text{the}^i; \{\text{the}^i; \text{man}^k\}\}$ . Given that Copy just replicates the targeted material and does not alter distinctiveness markings, it is possible to determine, at any point in a given derivation, whether two terms with identical sets of features are to be interpreted as copies or distinct constituents. (Nunes 2004: 22-23)

Such an implementation, besides violating the *Inclusiveness Condition* (Chomsky 1995), does not provide a principled explanation of what Non-Distinctiveness is (cf. Leung 2007, Neeleman & van de Koot 2010): it is only a marking mechanism; a way of implementing Non-Distinctiveness and not a real explanation of what that is. Thus, the Sameness Challenge is not properly resolved under Nunes (1995, 2004) theory.

#### 2.2. THE UNIQUENESS CHALLENGE

We may summarize this problem by asking why (5b) is unacceptable.

- (5) a. John was kissed ~~John~~.  
b. \*John was kissed John.

According to Nunes, the fact that only one copy is pronounced follows from the *Linear Correspondence Axiom* (LCA) of Kayne (1994). LCA imposes an antisymmetry requirement on syntactic structure and linear orderings:

- (6) *Antisymmetry*  
If X precedes Y, then Y cannot precede X.

Since two (or more) copies of a constituent are “the same”, a PF representation like (5b) would not comply with the antisymmetry condition (e.g., *John* precedes *was* and *was* precedes *John*). Thus, to linearize this kind of representations it is necessary to reduce all movement chains to only one link. The mechanism in charge to do so is called *Chain Reduction*.

- (7) *Chain Reduction* (Nunes 2004: 27)  
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.

Since this definition encodes a very explicit economy condition on the application of Chain Reduction (“delete the minimal number of constituents”), it is implied that (7) will not be applied when it is not necessary for convergence. Such a scenario will be discussed in section 4.

### 2.3. THE RANK CHALLENGE

Again, the problem may be summarized in a pair. Why is (8b) unacceptable?

- (8) a. John was kissed ~~John~~.  
b. \*~~John~~ was kissed John.

An obvious observation first: in order to explain why the higher copy is usually pronounced, we need to assume that there is some kind of **asymmetry** between the copies in a chain. It is very simple: if all copies were *exactly the same*, there would be no way to explain why one of them, the higher copy in the chain, is systematically selected to be pronounced.

In this sense, despite the fact that there is a “sameness” relation between the copies in a non-trivial chain, **their formal features may be different**. This is clear in (9), where the lower copy of John bears an uninterpretable Case feature (uC), whereas the higher copy has already checked that feature with the T head.

- (9)  $[_{TP} John_C [_{T'} was [_{VP} kissed John_{uC}]]]$

This representation is not legible at PF since, by definition, uninterpretable features (uFF) cause the crash of the derivation at the interfaces. Thus, it is necessary to assume that some kind of operation applies to (9) in order to eliminate all the uFF. This operation is defined in (10).

- (10) *FF-Elimination* (Nunes 2004: 31)  
Given the sequence of pairs  $\sigma = \langle (F, P)_1, (F, P)_2, \dots, (F, P)_n \rangle$  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.

By assuming (10) it is possible to explain the contrast between (8a) and (8b). If, as in (8a), the lower copy of *John* is deleted, then it is not necessary to apply FF-Elimination since Chain Reduction has already deleted the problematic uC feature; on the contrary, if the higher copy of *John* is deleted, as in (8b), then it will be necessary to apply FF-Elimination on the uC of the lower copy to make the representation legible at PF. Schematically:

- (11) a.  $[_{TP} John_C [_{T'} was [_{VP} kissed John_{uC}]]]$  ONLY REQUIRES CHAIN REDUCTION  
b.  $[_{TP} John_{uC} [_{T'} was [_{VP} kissed John_{uC}]]]$  ALSO REQUIRES FF-ELIMINATION

Given that the derivation in (11a) is shorter and involves fewer operations than (11b), the later is blocked by a (global) economy consideration. However, it is also predicted that when a shorter derivation leads to a crash, the less economical derivation will be acceptable. Such a scenario will be discussed in section 4.

If derivational economy is assumed (as in most versions of the Minimalist Program), grammatical operations may only take place if they prevent the syntactic derivation from crashing. The classic principle implementing this idea is the *Last Resort* condition:

- (12) *Last Resort* (Collins 1997: 9)  
An operation OP involving  $\alpha$  may apply only if some property of  $\alpha$  is satisfied.

Uninterpretable features are a particular implementation of this idea: they are properties of syntactic objects that need to be satisfied to make the syntactic representation legible at the interfaces. This is, at least, the way they work in Chomsky’s (1995) system, where the features of the elements in a chain are supposed to be some kind of unit.

- (13) *Uninterpretable feature checking in Chomsky’s (1995) system*  
a.  $[_{TP} T [_{VP} DP_{uF} [_{V'} \dots]]]$  UNINTERPRETABLE REPRESENTATION  
b.  $[_{TP} DP_{uF} [_{T'} T [_{VP} DP_{uF} [_{V'} \dots]]]]$  INTERPRETABLE REPRESENTATION

The problem with a representation as (13b) is that both copies are exactly the same regarding their features, so no asymmetry persists to explain why the higher copy should be pronounced.

Nunes's system, on the other hand, maintains certain asymmetry between the copies in a chain: only new copies check their uFF, while the lower ones preserve them. This implementation raises a problem for the very postulation of uninterpretable features. Consider the following derivation.

- (14) *Uninterpretable feature checking in Nunes' (1995, 2004) system*
- |    |   |                                |
|----|---|--------------------------------|
| a. | $[_{TP} T [_{VP} DP_{uF} [_{V'} \dots]]]$                 | UNINTERPRETABLE REPRESENTATION |
| b. | $[_{TP} DP_{uF} [_{T'} T [_{VP} DP_{uF} [_{V'} \dots]]]]$ | UNINTERPRETABLE REPRESENTATION |

If the application of movement operations is vacuous regarding the interpretability of the original occurrence of the DP, then it cannot be maintained that the operation depicted in (14) complies with the Last Resort condition. And, as said, since uFF are part of the technical implementation of Last Resort, their postulation in this system becomes gratuitous<sup>2</sup>.

As a précis, *the form of the argument* is the following: (i) there must be some asymmetry between copies to explain which one is pronounced; (ii) since they are "copies", such an asymmetry must proceed from the syntactic engagement of each copy with the main structure they were merged with; (iii) if this asymmetry is implemented in terms of uFF, we have a problem: uFF need to be deleted by syntactic mechanisms to comply with the Last Resort Condition; (iv) but if the uFF are deleted, there is no asymmetry between copies<sup>3</sup>.

### 3. An inclusion-based approach to chain formation in PF

The system I am going to sketch here is based on four main assumptions regarding the nature of syntactic computation.

- (15) *Assumption A*  
Syntax is about sets.

Following an important tradition in generative grammar (e.g., Chomsky 1995), I assume that lexical items are sets of features and phrases are sets of lexical items. This is, of course, a very usual idea.

<sup>2</sup> This criticism is only valid if *Local Economy* is assumed (cf. Collins 1997), which is the standard case after Johnson & Lappin's (1997) critique of *Global Economy* conditions.

<sup>3</sup> Personally, I think there is some tension between a principled implementation of Copy Theory and uFF.

- (16) *Assumption B*  
Late insertion: phonological information is introduced at PF.

The idea is that syntactic categories are purely abstract, having no phonological content (cf. Halle & Marantz 1993). The phonological matrixes are introduced post-syntactically by an operation called *Vocabulary Insertion* (cf. Embick & Noyer 2004). *Late Insertion* is a hypothesis independently motivated by syncretic (cf. Bonet 1991) and allomorphic (cf. Embick 2010) patterns.

Thus, a first big difference with Nunes' system arises. In Nunes' system, the phonological features of syntactic terminals are present throughout the entire syntactic derivation, so the lower links in a chain should be *deleted* at PF. Significantly, Nunes' answers for the Uniqueness and Rank Challenges rely on economy conditions on deletion operations.

In the system I am presenting here, syntactic terminals are phonologically "empty", so my answers for the Uniqueness and Rank Challenges cannot be grounded on economy considerations on deletion.

- (17) *Assumption C*  
There are no uFF in *Goals* (only in *Probes*).

As said, uFF are, at least, problematic to postulate the asymmetries between the links in a chain. For the moment, let's assume that movement operations are triggered by a functional head  $X^0$  attracting an YP to its specifier position and not by a specific requirement of YP (cf. Preminger 2011).

- (18) *Assumption D*  
If an XP enters in an *Agree* relation with a (non-defective) head  $H^0$ ,  $H^0$  assigns a feature [H] to XP.

So, for example, if a DP agrees with an inflected T head, the DP will be assigned a T-feature (nominative Case, following Pesetsky & Torrego 2001); or if a wh-phrase agrees with an interrogative complementizer  $C_{INT}$ , the wh-phrase will be assigned a Q-feature<sup>4</sup>.

Let's see how this system works.

Suppose an scenario where a non-defective *Probe*  $Y^0$  with an EPP-feature enters in an *Agree* relation with a *Goal*  $\alpha_{\{\dots\}}$ , where  $\{\dots\}$  is used to indicate the set of features of  $\alpha$ .

- (19)  $[_{YP} Y^0_{EPP} [_{XP} \dots \alpha_{\{\dots\}}]] \rightarrow [_{YP} \alpha_{\{Y, \dots\}} [_{Y'} Y^0_{EPP} [_{XP} \dots \alpha_{\{\dots\}}]]]$

<sup>4</sup> I am assuming that A'-dependencies also involve Probe-Goal mechanisms. Thus, intermediate/non-wh complementizers will count as "defective" regarding wh-phrases.

There is an *inclusion* relation between both copies of  $\alpha$ : the set  $\{\dots\}$  is a subset of  $\{Y, \dots\}$  ( $\alpha_{\{\dots\}} \subseteq \alpha_{\{Y, \dots\}}$ ). This relation between copies is going to arise systematically every time a new copy is generated<sup>5</sup>.

Since the inclusion relation between copies is systematic, it is possible to use it as a condition on chain formation:

(20) *Conditions on Chain Formation*

Two constituents  $\alpha$  and  $\beta$  are links in the same chain iff

- a.  $\alpha$  is a superset of  $\beta$ ;
- b.  $\alpha$  c-commands  $\beta$ ;
- c. There is no syntactic object  $\gamma$  between  $\alpha$  y  $\beta$  such as (i)  $\beta$  is a superset of  $\gamma$  and (ii)  $\gamma$  is not a superset of  $\alpha$ .

The conditions in (20) predict that a chain is formed every time the features of a lower constituent are included in a higher constituent. Also, these conditions make no use of markings to indicate when two constituents are copies, so they comply with Inclusiveness.

The type of “sameness” in (20a) is simpler than other potential definitions of “sameness”.

(21) *Kinds of “sameness”*

- a. Equality: (CH = (A, B), iff  $A = B$ )<sup>6</sup>
- b. Non-Distinctiveness: (CH = (A, B), iff  $A^i$  and  $B^i$ )
- c. Inclusion: (CH = (A, B) iff  $B \subseteq A$ )

Remember: lexical items and phrases are sets, so they are subject to the following definition of equality:

(22) *Set equality*

A set  $A$  is equal to a set  $B$  ( $A = B$ ) iff (i)  $A$  is a subset of  $B$  ( $A \subseteq B$ ), and (ii)  $B$  is a subset of  $A$  ( $B \subseteq A$ ).

Thus, even if a principled explanation for the “sameness” types in (21a) and (21b) is proposed, the option involving inclusion would be preferable due its simplicity.

<sup>5</sup> Even when the new copy does not receive an additional feature, it will be a superset of the original copy (e.g.,  $\alpha_{\{\dots\}} \subseteq \alpha_{\{Y, \dots\}}$ ). The notions of subset/superset assumed here must not be confused with the more complex notions of *proper subset/superset* (e.g.,  $A$  is a Proper Subset of  $B$  if  $A \subseteq B$  and  $A \neq B$ ).

<sup>6</sup> Remember: this definition is not going to work since some asymmetry between copies is necessary to decide which one is going to be pronounced.

Once the Sameness Challenge has been surpassed, we should focus on the Uniqueness Challenge and the Rank Challenge. For expository purposes, I will use the following pattern.

- (23) a. John was kissed ~~John~~
- b. \*~~John~~ was kissed John
- c. \*John was kissed John
- d. \*~~John~~ was kissed ~~John~~

In a Late Insertion system, the phonological realization of the syntactic terminals depends on the application of Vocabulary Insertion. Thus, an economy criterion on its application may explain the unacceptability of (23c), answering the question posed by the Uniqueness Challenge<sup>7</sup>.

We still need to explain the rest of the pattern:

- (24) a. John was kissed ~~John~~
- b. \*~~John~~ was kissed John
- d. \*~~John~~ was kissed ~~John~~

A simple way to derive such a pattern is by appealing to general conditions on recoverability of deletion, “which we may assume to exist though they are not understood in detail” (Chomsky 1977: 86). An operative definition of these conditions (based in Chomsky’s 1964 original postulation<sup>8</sup>) is proposed in (25):

(25) *Condition on Recoverability of Deletion*

An element may be deleted (i.e., not pronounced) if it is totally determined by a structurally related syntactic constituent.

Thus, assuming (25) it is possible to derive the necessity of pronouncing at least one copy (which derives the unacceptability of (23d)) and it is predicted that the higher copy in the chain (the one that engaged in more Agree relations) is the one that should receive phonological representation<sup>9</sup>. This theoretical result is captured in the following principle:

<sup>7</sup> There are, at least, two other ways to surpass the Uniqueness Challenge. The first one is by assuming that linear orderings need to comply with an antisymmetry requirement, just as Nunes (1995, 2004) did. The other one is, I think, theoretically more appealing: it consists on assuming that Vocabulary Insertion is an operation that takes (trivial and non-trivial) chains as inputs. This second option, however, requires further elaboration.

<sup>8</sup> “A transformation can delete an element only if this element is the designated representative of a category, or if the structural condition that defines this transformation states that the deleted element is structurally identical to another element of the transformed string” (Chomsky 1964: 41).

<sup>9</sup> In a Late Insertion system, this result is necessary given the logic of the *Subset Principle* (Halle 1997). The Subset Principle presupposes that “the terminal nodes that are the sites for insertion are

- (26) *Pronounce the Superset*  
Given a Chain CH, apply Vocabulary Insertion to the link which is the superset of all the remaining links.

#### 4. Deriving the exceptions to the Uniqueness and Rank Challenges

Nunes discusses two very important exceptions to the challenges stated in (3b) and (3c). First, again Uniqueness, in some languages it is possible to pronounce two (or even more) links of the same chain.

- (27) *Romani (McDaniel 1986)*  
**Kas** misline **kas** o Demiri dikhla?  
Who think who Demir saw  
'Who do you think Demir saw?'
- (28) *German (McDaniel 1986)*  
**Wen** glaubt Hans **wen** Jakob gesehen hat?  
Who thinks Hans who Jakob seen has  
'Who does Hans think Jakob saw?'
- (29) *Dutch (Barbiers et al. 2010)*  
**Wie** denk je **wie** ik gezien heb?  
Who think you who I seen have  
'Who do you think I have seen?'

Second, against Rank, there are some cases in which it is necessary to spell-out a copy which is not the higher link in a chain. According to Bošković's (2000, 2002) analysis, this is attested in Serbo-Croatian, a multiple *wh*-fronting language in which two adjacent homophone interrogative pronouns cannot be pronounced.

- (30) a. Ko šta kupuje?  
Who what buys  
b. \*Ko kupuje šta  
Who buys what  
'Who buys what?'
- (31) a. \*Šta šta uslovljava?  
What what conditions  
b. Šta uslovljava šta?  
what conditions what  
'What conditions what?'

fully specified; that is to say, they contain a full complement of syntactico-semantic features" (Embick & Noyer 2004: 299).

- (32) a. Šta neprestano šta uslovljava?  
What constantly what conditions  
b. \*Šta neprestano uslovljava šta?  
What constantly conditions what  
'What constantly conditions what?'

These patterns are very hard to explain in an approach to Copy Theory that, for example, assumes that the sensorimotor interface pronounces elements in the positions where their relevant morphophonetic features are checked (cf. Richards 2001, Stroik 2009), since in both cases constituents that did not check their features by syntactic means are pronounced.

Nunes (2004) argues that the first kind of pattern may be explained if it is assumed that an intermediate copy of the *wh*-pronoun may be morphologically reanalyzed as part of some functional head (e.g., a complementizer).

- (33) [<sub>CP</sub> Wen<sup>i</sup> ... [<sub>CP</sub> [<sub>C°</sub> wen<sup>i</sup> C<sup>0</sup>] [<sub>TP</sub> ... wen<sup>i</sup> ... ]]

Therefore, assuming Chomsky's (1995) proposal that the LCA does not apply word-internally, the phonetic realization of the intermediate copy of the *wh*-pronoun follows from economy considerations: since it will not violate the antisymmetry condition (because it is "invisible" to the LCA), it is not necessary to erase it.

This analysis makes two very strong predictions regarding the nature of the multiple copies: (i) only cyclic movement-based copies in "intermediate positions" may produce the kind of configuration that is necessary to this kind of reanalysis, so it is predicted that there is no multiple copy pronunciation involving the tail of the chain (cf. (34) vs. (35)); and (ii) since Nunes assumes that the reanalysis is based in an application of the morphological operation *Fusion*, it is predicted that there is no multiple copy pronunciation of full phrases (36).

- (34) *German (Fanselow and Mahajan 1995)*  
**Wen** denkst Du **wen** sie meint **wen** Harald liebt?  
who think you who she believes who Harald loves  
'Who do you think that she believes that Harald loves?'
- (35) *German (Nunes 2004)*  
\***Wen** glaubt Hans **wen** Jakob **wen** gesehen hat?  
whom thinks Hans whom Jakob whom seen has  
'Who does Hans think Jakob saw?'
- (36) *German (Nunes 2004)*  
\***Wessen Buch** glaubst du **wessen Buch** Hans liest?  
whose book think you whose book Hans reads

'Whose book do you think Hans is reading?'

A similar kind of analysis may be implemented in the inclusion-based system presented here. The following is the syntactic representation I am assuming for (28). Here the wh-pronoun entered in an Agree relation with the interrogative complementizer C<sub>Q</sub> and, given the assumption in (18), it receives a Q-feature.

(37) [CP **wen**<sub>{Q, Acc, D, φ}</sub> [C' C<sub>Q</sub> ... [CP **wen**<sub>{Acc, D, φ}</sub> ... [TP ... [VP V **wen**<sub>{Acc, D, φ}</sub>]]]]]

Following Nunes (2004), I will assume that the morphological reanalysis between the wh-pronoun and C involves an application of the operation *Fusion*, although I am assuming a somewhat different definition.

(38) *Fusion* (Embick 2010: 78)  
 $[_x \alpha] \frown [_y \beta] \rightarrow [_{x/y} \alpha, \beta]$   
 where  $\alpha$  and  $\beta$  are features of X and Y.

When subordinate C and *wen* fuse, the result is a syntactic terminal with a set of features containing both the features of *wen* and C:

(39) [CP **wen**<sub>{Q, Acc, D, φ}</sub> [C' C<sub>Q</sub> ... [CP **#C + wen**<sub>#{C, Acc, D, φ}</sub> [TP ... [VP V **wen**<sub>{Acc, D, φ}</sub>]]]]]

According to the Conditions on Chain Formation in (20), there are two chains in (36). And, as normal, the head of each chain is pronounced in both cases.

(40) a. CH<sub>1</sub> = (**wen**<sub>{Q, Acc, D, φ}</sub>)  
 b. CH<sub>2</sub> = (**#C + wen**<sub>#{C, Acc, D, φ}</sub>, **wen**<sub>{Acc, D, φ}</sub>)

Thus, in the system I am presenting here, multiple copy phenomena imply multiple chains at PF.

Regarding low copy pronunciation data in (30) to (32), Nunes suggests (following a proposal in Franks (1998)) that Chain Reduction applies on the higher copy in a non-trivial chain when it is necessary to satisfy some PF language specific requirement; in this case, *Antihomophony*. This idea can be translated into the inclusion-based system by stating Antihomophony as a language specific condition on Vocabulary Insertion.

(41) *Antihomophony*  
 Vocabulary Insertion cannot apply two consecutive times for the same Lexical Item.

So, assuming that Vocabulary Insertion applies “from left to right”, (41) allows explaining the Serbo-Croatian pattern.

(42) Šta<sup>i</sup> šta<sup>i</sup> uslovljava šta<sup>i</sup>

## 5. Conclusions

- The standard approach to the three Challenges of Copy Theory (Nunes 1995, 2004) suffers some important drawbacks with the Sameness Challenge and the Rank Challenge.
- These drawbacks are avoided by assuming that the “sameness” relation underlying chains follows from an *inclusion* relation between syntactic objects.
- The system based on inclusion also derives the exceptions to the general patterns (namely, multiple copy phenomena and low copy pronunciation).

## References

- Barbiers, S., Koenenman, O., & Lekakou, M. 2010. Syntactic doubling and the structure of wh-chains. *Journal of Linguistics* 46: 1-46 || Bonet, E. 1991. *Morphology after syntax: Pronominal clitics in Romance*. Doctoral Dissertation, MIT || Bošković, Ž. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33: 351-383 || Chomsky, N. 1964. *Current Issues in Linguistic Theory*. Mouton: The Hague || Chomsky, N. 1977. On Wh-movement. In P. Culicover, T. Wasow & A. Akmajian (eds), *Formal Syntax*. New York: Academic Press || Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris || Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale & S.J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge: MIT Press || Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press || Chomsky, N. 2001. *Derivation by Phase*. Michael Kenstowicz (ed.), *Ken Hale. A Life in Language*. Cambridge, MA: MIT Press || Collins, C. 1997. *Local Economy*. Cambridge: MIT Press || Embick, D. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press || Embick, D. & Noyer, R. 2007. Distributed Morphology and the Syntax-Morphology Interface. In G. Ramchand and C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*. Oxford: OUP || Fanselow, G., & Mahajan, A. 1995. Partial movement and successive cyclicity. In U. Lutz and G. Müller (eds.), *Papers on wh-scope marking*. University of Stuttgart and University of Tübingen || Franks, S. 1998. Clitics in Slavic. Paper presented at the *Comparative Morphosyntax Workshop*, Bloomington, Ind. || Halle, M. 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics* 30: 425-449 || Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In K. Hale & S.J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge: MIT Press || Johnson, D., & Lappin, S. 1997. A critique of the Minimalist Program. *Linguistics and Philosophy* 20: 273-333 || Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press || Leung, T. 2007. *Syntactic derivation and the theory of matching contextual features*. Doctoral Dissertation, University of Southern California || McDaniel, D. 1986. *Conditions on wh-chains*. Doctoral Dissertation, City University of New York || Neeleman, A., & Van De Koot, H. 2010. A local encoding of syntactic dependencies and its consequences for the theory of movement. *Syntax*, 13: 331-372 || Nunes, J. 1995. *The copy theory of movement and the linearization of chains in the minimalist program*. Doctoral Dissertation, University of Maryland || Nunes, J. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press || Pesetsky, D. & Torrego, E. 2001. T-to-C movement: causes and consequences. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, Mass.: MIT Press || Preminger, O. 2011. *Agreement as a fallible operation*. Doctoral Dissertation, MIT || Richards, N. 2001. *Movement in Language*. Oxford: OUP || Stroik, T. 2009. *Locality in Minimalist Syntax*. Cambridge: MIT Press.