Connectivity in Markovian dependencies

Ángel J. Gallego

This paper explores the properties (and limits) of Merge, arguing, with Uriagereka (2005, forthcoming), that it is a natural reflection of the so-called "Chomsky Hierarchy". Under that scenario, I focus on what the mechanisms current syntactic theory provides to study adjuncts are, the conclusion being that these dependents manifest two basic structural realizations: *bona fide* adjunction (however this must ultimately be captured –plausibly, through Chomsky's 2004 pair-Merge) and modification from a specifier position in a cascade structure. As will be shown, the properties of both structural possibilities cannot be reduced nor unified, so different empirical predictions are made in each case.

1. Bare phrase structure

Since Chomsky (1995), X-bar theory algorithms like those in (1) are eliminated from the system due to economy and theory internal reasons.[1]

(1) a. XP → SPEC X'
b. X' → Xº COM

With the X-bar schemata gone, minimalism resorts to a Bare Phrase Structure (BPS) with the basic (though potentially unbounded) computational operation of Merge, which takes two syntactic objects (SO)[2] (say, α and β) and yields a bigger one, K, as depicted in (2).

(2) K = {α, β}

What is K? In the formulation of Chomsky (1995), K is supposed to adopt the form of (3), where γ encodes the category (or semantic type) of the resulting SO.

(3) K = {γ, {α, β}}

[1] See Chomsky (1970) for the original formulation of X-bar theory. See Chametzky (2000) and Fukui (2001) for extensive review and discussion.

[2] I assume that both Lexical Items (LIs) and phrases (i.e., combinations of LIs) count as SOs.

Proceedings of ConSOLE XIV, 2007, 73-98
http://www.sole.leidenuniv.nl

Chomsky (1995) calls γ 'label'; at present there is no consensus as to whether BPS must make use of labels or not (see Boeckx 2002 for discussion)[3], and although Chomsky (p.c.) points out that BPS is, in itself, independent of labels, he has also noted that these seem to be needed for computational reasons. In his own words:

> "Each syntactic object generated contains information relevant for further computation. Optimally, that will be captured entirely in a single designated element, which should furthermore be identifiable with minimal search: its label, the element taken to be 'projected' in X-bar-theoretic systems. The label, which will invariably be a lexical item introduced by Merge, should be the sole probe for operations internal to the syntactic object, and the only element visible for further computation […] Note that labels, or some counterpart, are the minimum of what is required, on the very weak assumption that at least some information about a syntactic object is needed for further computation, both for search within it and for its external role." [Chomsky (2005a:14)]

By and far Collins (2002) is the most remarkable phrase structure proposal without labels, but although his system is label-free indeed, it can still be said to contain a residue of these symbols, the Locus. What is the Locus? Collins (2002) argues that it is an LI that needs to be 'saturated', the element that 'selects', the one which is 'active' during the computation. To understand this notion better, consider the Locus Principle in (4), taken from Collins (2002).

(4) **The Locus Principle**
Let X be a lexical item that has one or more probe/selectors. Suppose X is chosen from the lexical array and introduced into the derivation. Then the probe/selectors of X must be satisfied before any new unsaturated lexical items are chosen from the lexical array. Let us call X the Locus of the derivation. [Collins (2002:46)]

The formulation in (4) correctly captures the dynamic nature of Collins' (2002) Locus, contrary to the representational status of labels: at every derivational stage there is only one Locus, while there can be many labels.

Before going on, some technical aspects of (4) should be clarified. This is the case of notions like Goal and Probe, which belong to the operation Chomsky (2000) associates to Case checking: Agree. Roughly put, Agree establishes a (long-distance) dependency between two syntactic objects for feature valuation purposes. This can be seen in (5).

(5) John loves Mary.

Under fairly standard assumptions, the DPs *John* and *Mary* are supposed to check their Case feature against the functional categories T and *v**.[4] The process, in essence, works like this: T and *v** have uninterpretable φ-features which –Chomsky argues– enter the syntactic component unvalued, and hence act as a seeker (a Probe) looking for an element with which

[3] To the best of my knowledge, Moro (2000) is the first BSP-based account in which SOs are taken to be label-free. In particular, Moro (2000) provides robust empirical evidence that a variety of Small Clauses (those which he dubs 'Bare Small Clauses') do not project label. The same basic idea has been applied to adjuncts by Chametzky (2000), Hornstein et al. (2005), and Uriagereka (2003).

[4] See Chomsky (2005b) for slightly modified ideas on Case.

to Match and value its features (a Goal). As a result of that matching, the Probe gets its features valued and, in exchange, assigns Case to the Goal.[5]

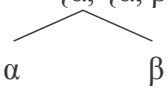
To recap so far, it is easy to see a connection holding among Probe, Label, and Locus: these are just different names for the element which drives the computation. There is, though, a minor qualification to this parallelism: on derivational grounds, only Probe and Locus share the relevant 'viral' property (in Uriagereka's 1998 sense) of triggering a scanning procedure.

In this paper I would like to defend the role of labels –much in the sense of Boeckx (2002) and Hornstein (2005)-, for which some arguments will be provided.[6] Meanwhile, let us go back to the basic operation of Merge, and, more particularly, to one aspect of its output: the label. As is well known, Chomsky (1995:244) contemplates the possibilities in (6).

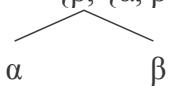
(6) The label γ is...
- a. The intersection of α and β (i.e., $\alpha \cap \beta$).
- b. The union of α and β (i.e., $\alpha \cup \beta$).
- c. One or the other of α, β.

Arguably because he assumes a decompositional approach to syntactic categories in terms of [±N] [±V] features (see Chomsky 1970, 1981), Chomsky (1995) rejects both (6a) and (6b), taking (6c) to be the correct outcome.[7] Merger of α and β will then produce one of the two sets in (7), depending on what the label turns out to be:

(7) a. {α, {α, β}}



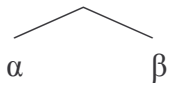
b. {β, {α, β}}



Note that first-Merge, as in (7), gives us the basic relation of "set-membership" (or "immediate containment") and, applied twice, the relations "contain" and "term-of". Adding "sisterhood" to the picture, "c-command" obtains.

From BPS it is also possible to derive the bar-levels of previous theories: Xº, X', and X''. Any LI is, unless qualified, both X^{max} (a phrase) and X^{min} (an LI or head): if it does not project, it is an X^{max}, otherwise it is an X^{min}. As for X', it is eliminated from the system.[8] But note that there is a catch: for the system to be able to make these distinctions, labels (or some counterpart) are needed. We cannot know whether β is an LI or an XP in (8) without labels:

(8)



[5] Note that 'valuation' is the only thing syntax cares about; 'interpretability' (or its lack) is an interface issue, irrelevant for computation as such. See Pesetsky & Torrego (2004) for discussion.

[6] See Boeckx (2002) for a more detailed discussion of Collin's (2002) proposal.

[7] The evidence in favor of (6c) is not strong. Consider the union case: it is not obvious that if N and V Merge, there is a problem, since, after all, we can have sets like (i), as noted by Norbert Hornstein (p.c.):

(i) {+1, -1}

The question that (i) (and the merger of N and V, for that matter) raises is whether a set with 1 and -1 violates the Law of the Excluded Middle: A is B or A is not B. Formally:

(ii) $A \vee \neg A$

[8] See Chomsky (1995) for discussion about invisibility of X'. Under strongly derivational models, like Epstein et al.'s (1998), X' could be taken as a residual X^{max}.

Can the choice of the label be predicted? Chomsky has gone back and forth in this respect. In the 1995 and 2005b formulations, he seems to suggest that the label cannot be decided a priori: any SO undergoing Merge can project, but if the wrong choice is made, there will be a crash at the interfaces. Chomsky (2000), though, presents in passing a different view whereby Merge has an Agree-like nature: either α or β acts as a Probe, and projects. A compatible view is put forward by Boeckx (2002), who regards labeling as a feature sharing operation (see Frampton & Gutmann 2000 and Pesetsky & Torrego 2004); in particular, Boeckx (2002), aiming at unifying Merge and Move, proposes that whenever two LI undergo Merge, one of their features must be matched and 'percolate up', as a result of intersection:

(9) The label K of {α, β} = The feature F shared in grouping {α, β}.
[from Boeckx (2002:21)]

As Boeckx (2002) notes, the mechanics of (9) are similar (if not identical) to both Attract-F and Agree of previous minimalism formulations. The problem about labelling noted by Chomsky (1995) does not arise if categorial features (e.g., [±N] and [±V]) are replaced by φ-features (see Chomsky 2001 and Marantz 1997).[9]

Whatever the particular implementation, let us suppose Merge provides a label. This also allows us to derive the most local relation: the Head-to-Head one. Depending on the result of merging α and β, one of the heads becomes what we call Complement, as we see in (10b):

(10) a. α (Head) β (Head) ⇒ b. {α, {α, β}}
α (Head) β (Complement)

Is there any other relevant notion? Syntactic theory has devoted much attention to specifiers (SPECs). Under BPS, these result from all non-first-Merge labelled operations, as shown in (11):

(11)
α
δ (SPEC) α
λ (SPEC) α
α (Head) β (Complement)

As the reader may easily calculate, there is only one head and one complement per projection, but a potentially unbounded number of SPECs. That is an asymmetry. Also, note that for us to know that something is a SPEC, we must have labels: if no label projects when α and β Merge in (11), we cannot know whether λ is a SPEC.

[9] This merely scratches the surface of Boeckx's (2002) technical implementation. See Gallego (2007) for more detailed discussion of Boeckx's (2002) proposal.

On the face of it, Chomsky (2004) argues that SPEC-Head relations lack conceptual motivation, and they should actually reduce to Head-Head ones as well: SPECs are always created by a feature of a head.[10]

Let us explore SPECs in more detail, assessing what types of SPECs we can have in the first place. As far as I know –and putting s-selection issues aside–, SPECs can be created by two varieties of Merge, what Chomsky (2004) calls external-Merge and internal-Merge (i.e., Move). The first type of Merge concatenates two completely independent SOs, taken from the Lexicon (12a), whereas in internal Merge one of the SOs is generated within the other (12b):

(12) a. $\{\alpha, \{\alpha, \beta\}\}$

α (Head) β (Complement)

b. $\{\alpha, \{\beta, \{\alpha, \{\alpha, \beta\}\}\}\}$

β (SPEC) $\{\alpha, \{\alpha, \beta\}\}$

α (Head) β (Complement)

Notice that, as described in (12b), internal-Merge displays one important property: it satisfies Chomsky's (1993) Extension Condition.[11]

Chomsky (1995) makes a subtle distinction that concerns how SPECs and labels interact: details aside, he argues that SPECs obtained by internal-Merge cannot project upon merger with their target; on the contrary, when SPECs are created by external-Merge, the label cannot be decided if the SOs undergoing Merge have the same structural status (i.e., if they are both Xs or XPs). Let us consider these two claims step by step. The details of the first one go back to the discussion in Chomsky (1995:4.4.2) and it surely has to do with both Donati's (2006) and Hornstein & Uriagereka's (2002) proposals, which I cannot go into here; at any rate, what matters: why does Chomsky (1995) claim that SPECs cannot project when they are created by internal-Merge? He adduces two problems[12]: first, the would-be checking configuration is not legitimate, and, second, the resulting chain is not uniform. We can dismiss these problems at once if we adopt a strongly derivational model and if checking configurations are dispensed with (see Chomsky 2000). Consider now Chomsky's (1995, 2005b) second claim: {X(P), Y(P)} structures created by external-Merge are label-free. Chomsky (2005b) reassesses this very point, mentioning Moro's (2000) findings on Small Clauses in order to argue that these and the merger of external arguments create "unstable structures" where no label can be decided.[13,14] This claim is interesting, and it seems to be empirically well motivated (see Moro

[10] See Starke (2004), who eliminates SPECs by claiming that these are to be 'recycled' as heads.

[11] This view is not shared by sideward-based approaches to Movement (see Nunes 2004). Actually, under Nunes' (2004) proposal, it makes no sense to speak about "internal-Merge", for Movement can occur among independent phrase markers (by means of massive generalized transformations). There are, I think, two potential drawbacks to Nunes' (2004) theory: first, it must assume an operation of Copy, independent from Merge; and, second, it blurs Chomsky's conceptual distinction between external and internal Merge. The first issue predicts that Copy could in principle operate on its own, without Merge, but this is not the case: as far as I know, Copy and Merge always work together. As for the second problem, it is related to the duality of semantics that distinguishes theta-theory from scope/discourse-related notions (i.e., presupposition, focus, specificity, etc.).

[12] See Chomsky (1995:4.4.2. & 4.7.2.)

[13] Although there may be some truth to it, I think it is hasty to speak about instability here. In Moro (2000), instability is defined in very accurate terms: the structures he explores create a "point of symmetry", and thus violate Kayne's (1994) LCA, so movement must operate for linearization to be possible. Such an explanation, regardless of its adequacy, is principled.

[14] In Gallego (2006) I argue that the same holds in the case of expletives, direct objects of transitive verbs (under Mateu's 2002 system), and adjuncts.

2000 for ample evidence that some Small Clauses lack a label; as for the merger of external arguments, perhaps Chomsky's 2005b reasoning may shed some light on why something must escape the *v**P).[15] I leave the technical details of this complex discussion here, with no additional useful comment. In section 4 I will go back to this issue, arguing that the same logic about instability in {X(P), Y(P)} structures can be extended to adjuncts.

The goal of this section was to introduce the elementary technical notions of BPS which should be borne in mind in the remainder of this paper. The discussion has focused on the status of labels within current theorizing. As Boeckx (2002), Chomsky (2005a) and specially Hornstein (2005) note, there are grounds to believe that no theory of grammar can eliminate labels. Labels naturally embody the notion of "constituent" and are therefore essential for structure building processes.[16] Thus, as Boeckx (2002) wisely notes, it is difficult to see how syntactic processes like VP topicalization (see (13a)), wh-movement (see (13b)) or VP ellipsis (see (13c)) could take place without labels, for these operations need to locate the relevant maximal/XP chunks that are going to be targeted.[17]

(13) a. $[_{VP}$ Kiss Mary$]_i$, John did t_i
b. Which book$_i$ did John say Mary read t_i ?
c. John called Peter, and Mary did Susan$_i$ ~~call~~ t_i.

Virtually any operation targeting XPs needs to invoke labels, and the logic extends to Agree if φ-features are encoded in labels.[18] The same seems to hold in the PF and LF wings of the grammar: if something like Kayne's (1994) LCA is correct, then labels are needed to calculate linear order.[19] As for LF, labels may also be useful to indicate semantic types.

Stronger conclusions can be drawn from Hornstein (2005), who argues that labels turn a complex SO into an atomic unit[20], which can then be used for further 'concatenation' – actually, for Hornstein (2005), without labeling no hierarchical embedding would be possible. From the necessity of labels, Hornstein (2005) also derives the Extension Condition and endocentricity.

For the purposes of this section, we can stop here. I will go back to some questions raised by the necessity of labels in section 4, where I explore adjuncts.

2. The Chomsky Hierarchy

Up to this point, we have seen two varieties of Merge: external-Merge and internal-Merge. Technically, only the first variety is context free, being thus expressible by Phrase Structure Grammar algorithms like (14), where A is a non-terminal symbol, and γ stands for a string of terminal or non-terminal symbols.[21]

---

[15] See Alexiadou & Anagnostopoulou (2001).

[16] I disregard here any technical distinction between nodes and labels. See Chametzky (2000).

[17] Although along this paper I will tacitly assume that internal-Merge leaves a copy of the moved element, I will represent it by means of a trace. See section 3.

[18] This is explicitly stated by Chomsky (1995:268), who refers to the features of labels as "sublabels".

[19] See Richards (2001) for explorations of this view.

[20] See Uriagereka (2004) for a similar reasoning about the creation of complex SPECs.

[21] A Phrase Structure Grammar (PSG) has four components: 1) a terminal vocabulary (a Lexicon); 2) a non-terminal vocabulary (the non-terminal symbols); 3) a set of rules; and 4) a starting symbol which is member of 2 (typically, the sentence node S). See Uriagereka (forthcoming) for relevant discussion.

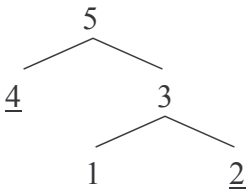
(14) $A \rightarrow \gamma$ [22]

The most remarkable property of context-free grammars is their ability to generate a system of rules such as (15a) or (15b), where the conversion process indicated in (14) affects A, regardless of the context in which it is found:

(15) a. PP → P NP
b. NP → N PP

Notice that, given two simple rules like (15a) and (15b), we can have a potentially unbounded structure, since part of the output of (15b) is the input of (15a), which ensues an endless loop. This is, incidentally, one of the problems of context-free grammars: they overgenerate. A second –and perhaps more important– problem concerns the impossibility of expressing a long-distance relation between SOs; that is to say, (14) cannot capture a possible dependency between 4 and 2 bypassing 1 in (16).

(16)



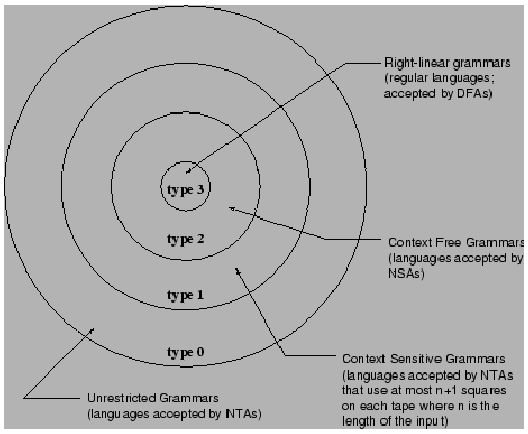
The context of all elements in (16) is formally well defined. However, with the limited apparatus of (14) we cannot express a connection between 4 and 2 ignoring 1, for this requires a context sensitive procedure.[23] For those familiar with the Minimalist Program, this type of (long-distance) dependency has been phrased in terms of Attract-F, Agree, and Move, and can be encoded by means of the algorithm in (17), where A is a terminal symbol, and α, β, and γ are sequences of terminal or non-terminal symbols:

(17) $\alpha\ A\ \beta \rightarrow \alpha\ \gamma\ \beta$

Context sensitivity follows, here, from the variables α and β, which define A's "context", allowing us to know whether this element can be replaced by γ, as the rule dictates. Both context free and context sensitive rules were studied and classified by Chomsky fifty years ago in what is known as the "Chomsky Hierarchy":

[22] This type of algorithm is also known as "rewriting rule", and is a normal device of PSGs. In plain terms, what a rule like (14) does is transforming the string of symbols to the left into that to the right.

[23] Such procedure predicts the possibility of so-called intervention effects (in the sense of Rizzi 1990): if the relevant property shared by 4 and 2 is also possessed by 1, then 1 will interfere. As Sergio Balari (p.c.) makes me note, if there is no minimality scenario (say, if 4, 2, and 1 do not share any property), the operation I have alluded to is context-free, not context-sensitive; for instance, non-multiple wh-movement (e.g. Who_i did John call t_i?) is context-free.



Types of Formal Languages

Type 0 → Recursive Languages
Sets of formal objects of any computational complexity

Type 1 → Context-sensitive Languages
Sets of sets of sequences of symbols (i.e., chains)

Type 2 → Context-free Languages
Sets of sequences of symbols (i.e., phrases)

Type 3 → Regular Languages
Sequences of symbols.

(18) The Chomsky Hierarchy (see Chomsky 1956)

Uriagereka (2005, forthcoming) reinterprets the Chomsky Hierarchy in terms of memory, as indicated in (19):

(19) a. Recursive grammars: unbounded memory.
b. Context-sensitive grammars: short term memory.
c. Context-free grammars: long term memory.
d. Regular grammars: no memory.

We have previously introduced the operations instantiating the types 1 and 2 of grammars (external-Merge and internal-Merge), but nothing has been said about types 0 and 3. Following Uriagereka (2005, forthcoming), I will assume that type 0 corresponds to the Turing Machine and is therefore out of the Faculty of Language.[24, 25] What about type 3? As a matter of simple logic, type 0 cannot run the same fate of type 3, since the Chomsky Hierarchy has an implicational nature: more complex levels presuppose simpler ones. (20) contains the rules that are typically associated to formal languages of type 3:

(20) a. $A \rightarrow \alpha\ \gamma$
b. $A \rightarrow \gamma\ \beta$

The basic property of these rules is their memory restrictions: being incapable of keeping track of complex associations (i.e., phrases), regular rules can only operate with immediately adjacent elements, as happens in the case of Markovian chains.[26]

[24] The formal proof goes back to Chomsky (1963). I thank Sergio Balari for discussing this point with me.

[25] The term "machine" is not intended in the contemporary sense. Alan Turing used "machine" in the same way we nowadays use the term "program",

[26] Or Markov Chain, in honor of the Russian mathematician Andrei Markov. These are sequences of symbols as such as 'X_1, X_2, X_3, X_4, ...' where whatever precedes a stage *s* depends on what happened in an immediately previous stage *s – 1*.

It is obvious that natural languages present phenomena which need more memory than that; what must be found out is whether they also present processes of the Markovian sort. The first claim can easily be tested in domains like thematic theory; consider (21):

(21) John ate the sandwich.

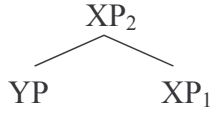
The question arises as to how we know that the relation between the DP *the sandwich* and the verb is different from that established between the DP *John* and the verb. Put another way: assuming that linear order is irrelevant, how do we choose between (22a) and (22b)?

(22) a. John ∩ ate ∩ the ∩ sandwich
b. [John [ate [the [sandwich]]]]

(22b) has everything (22a) does, plus something else. This "something else" is phrase structure. Importantly for my concerns, only by means of this "something else" we can encode the fact that the relation between objects and verbs is more intimate than the one between subjects and verbs (see Hale & Keyser 2002; Harley 2003; Marantz 1984,1997; and Tenny 1994), which determines that it is the object that is interpreted as a /Theme/ –the element that "measures out" the event, in Tenny's (1994) sense. Happily, in order to capture this more intimate relation between objects and verbs, the system must keep track of the fact that verbs and objects form a constituent which excludes the subject, and that requires phrase structure.

Let us now assess the second claim made above: given the implicational nature of the Chomsky Hierarchy, having formal languages of type 1 and 2 implies having formal languages of type 3 (i.e., Markovian processes). The ideal candidate to fulfill this type of operation is adjunction (see Chametzky 2000), which is formally represented as in (23):

(23)



In (23) YP is adjoined to XP. As can be seen, adjunction splits the XP category into two segments (XP_1 and XP_2), the hallmark of May's (1985) formulation, adopted by Chomsky (1986). Simplifying somewhat, (23) is designed to place the adjunct in a parallel dimension: by exploiting the category vs. segment category, adjuncts are immune to almost all the relevant structural relations one can think of (e.g., dominance, c-command, etc.). In Chomsky (2004) this "parallel plane" idea has been stated explicitly (see also Uriagereka 2003):

> "We have so far assumed only the free symmetrical operation Merge, yielding syntactic objects that are sets, all binary: call them *simple*. The relations that come "free" (contain, c-command, etc.) are defined on simple structures. But it is an empirical fact that there is also an asymmetric operation of adjunction, which takes two objects β and α and forms the ordered pair <β, α>, α adjoined to β. Set Merge and pair Merge are descendants of substitution and adjunction in earlier theories. Given the basic properties of adjunction, we might intuitively think of α as attached to β on a separate plane, with β retaining all its properties on the primary plane, the simple structure." [Chomsky (2004:117-118)]

As noted by many scholars (see Chametzky 2000; Boeckx 2003a, 2003b; Ernst 2002; Uriagereka 2003), Chomsky's (2004) words suggest that adjuncts have no syntax. This is surely consistent with most defining properties of adjuncts (e.g., they do not receive theta-roles, they do not check Case, they are islands, etc.) and with Chomsky's (2004) conceptual speculation that pair-Merge (the operation handling adjunction) exists by requirements imposed by the Intentional-Conceptual systems.

In the context of the present discussion, it is telling enough to remember that Chomsky (1995) claimed that adjuncts do not fit in the minimalist picture, for they do not seem to participate in any computational operation: an adjunction configuration is neither a thematic nor a checking one.[27]

> "The problems related to XP adjunction are perhaps a case in point: they may not really belong to the system we are discussing here as we keep closely to the first of the two courses just outlined, the one that is concerned with Last Resort movement driven by feature checking within the N → λ computation. It is within this core component of language that we find the striking properties highlighted by minimalist guidelines. It seems increasingly reasonable to distinguish this component of the language faculty." [Chomsky (1995:325)]

Let us now return to the technical details of adjunction. As we have seen, adjunction is formalized by means of ordered pairs within minimalism. This is ontologically worrisome, for we need to postulate an independent operation for adjunction; if possible (it is obviously desirable), adjunction should just resort to set-Merge, the aforementioned asymmetry being perhaps a consequence of the lack of label. Actually, Chomsky's (2004:117) formulation does not seem to me to be far from this, as he says that "an adjunction construction is plainly not the projection of a head: for NP-adjuncts, for example, the constituent structure appears to be something like [NP XP]". This possibility takes us back to Chomsky's (2005b) suggestion about unstable structures, which, interestingly, are compatible with what we see in any adjunction configuration within the *v**P: it is always the case that an XP (the adjunct) adjoins to YP (the VP).[28] The structure would therefore be something like [VP, XP] (by parity of reasoning with the nominal case pointed out by Chomsky 2004), so, formally, the whole structure would behave as a VP because the category which receives adjunction "retains all its properties". This is reinforced by the data in (24), where adding adjuncts does not modify the category/type of the syntactic object which gets modified; thus, as noted by Hornstein et al. (2005), given that perfective *have* subcategorizes for a perfective *–en* marked V, adding an adjunct like *quickly* or *in the yard* does not change its selectional requirements:

---

[27] Note that Chomsky's (1995, 2004) suggestions seem to point to the conclusion that adjuncts involve a more complex (parallel) structure. However, Norbert Hornstein (p.c.) and Wolfram Hinzen (p.c) make me note that there are grounds to believe that the opposite scenario is actually the correct one: arguments are more complex than adjuncts. Their observation makes sense in that it is adjuncts (not arguments, unless conjoined) that display a truly recursive nature; moreover, it is arguments (not adjuncts) that must invoke the notion of theta-role, undergoing a process of type-shifting which changes their status. True as this is, there is no a fortiori tension, for the idea that adjuncts are simpler than arguments is immediately captured by the analysis pursued here, where the former instantiate type 3 grammars, in Chomsky's (1956) terms.

[28] Under this formulation, Chomsky's (2004) original analysis is left basically intact, for adjuncts get transferred to the interfaces at the phase level, being in a configuration with the VP, the complement domain of the phase head *v**.

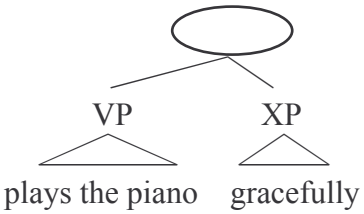
(24) a. Has/*is [$_{VP}$ eaten a bagel]
b. Has/*is [$_{VP}$ [$_{VP}$ [$_{VP}$ eaten a bagel] quickly] in the yard]

This very point is clearly put by Uriagereka (2003):

> "Labels are valid only if limited; an unbounded number of labels amounts to absence of generalization, or else a missed one […] For labels to be useful, they must be very few, so that computational tasks are simplified […] The result of the mere fact that adjunction to X is unbounded is clear: either we need an unbounded number of labels, or else each adjunction does not modify X's character. In both cases, labelling would be useless." [Uriagereka (2003:4)]

The idea is also sound within Boeckx's (2002) proposal: in adjunction, XP and YP do not obviously share any feature, so no label can obtain. My own interpretation of the process is as in (25), where I exclude the external argument from the picture.[29]

(25)



What happens if no label can be decided in (25)? Chomsky (2005b) suggests that something must raise, due to formal instability. A more interesting observation was (originally, to my knowledge) made by Cedric Boeckx: if no label is projected in (25), the system cannot know whether the VP or the XP is a (complex) SPEC, and then an island (see Huang 1982 and Uriagereka 1999). Notice that the latter point depends on the correlation between "being a (complex) SPEC" and "being an island" –Uriagereka's (1999) insight; but there are grounds to think otherwise, since, as noted by Uriagereka (1988) himself, postverbal subjects (which can be analyzed as bona fide SPECs) allow sub-extraction in some languages. Not everything is lost, luckily. We can still appeal to (25) in order to predict the extraction facts; as Norbert Hornstein (p.c.) notes, if something tries to move out of XP, a first logical step involves the creation of a SPEC with whatever the unlabeled node turns out to be, and right there we are stuck: we want to provide a SPEC-slot to the extracted element, but since there is no label, it is impossible.

Because it underscores the role of structure, a label-free analysis of adjuncts along these lines can account for some intriguing phenomena related to adjunction (which I will review in section 4) in an elegant fashion. In the preceding lines I have established a parallelism between the Chomsky Hierarchy and the three varieties of Merge entertained within BPS. Pair-Merge is undoubtedly the most interesting one, for it focuses on adjuncts –syntactic dependents showing a bunch of odd properties. I have also drawn a parallelism between Markovian syntax and adjuncts, but more needs to be said if we want it to be borne out, an issue I return to in section 4.

[29] In Gallego (2006, 2007) I brush up the implementation by taking Mateu's (2002) decompositional analysis of adverbs, which I apply to all adjuncts.

3. Connectivity: possibilities and mechanisms

This section is devoted to study some empirical facts concerning adjuncts in the light of what I will call "connectivity", a term I use to subsume any process that relates two structurally well defined positions in a given phrase marker. Many phenomena can be so defined (e.g., NPI licensing, scope dependencies, movement, and binding, to name but a few), and they tend to display an atypical behaviour when it comes to adjuncts. In this section I will restrict my attention to binding, going back to other kinds of connectivity in section 4.

To begin exploring binding, consider the sentence in (26):

(26) $[_{CP}\ [_{TP}$ John's sister$_i$ was $[_{vP}$ invited t_i to the party]]]

In pre-minimalist models, the displacement operation in (26) was assumed to leave a designated symbol in the base-position after the DP *John's sister* moved: a trace (which had to be further coindexed with and be c-commanded by the DP). This scenario changed in Chomsky (1993), where the copy theory of movement was revamped; basically put, the new idea (actually, old) was that movement leaves an exact copy of the moved element which is deleted at the phonological component due to economy reasons. The copy theory soon proved useful to account for some loose ends of binding, which from Chomsky (1993) on was taken to apply only at LF (the SEM component), allowing to dispense with the operation of reconstruction. So, in (27), the wh-phrase *what picture of herself* leaves a copy in the base position (occupied by the trace), allowing Principle-A to go through.

(27) $[_{CP}$ [What picture of herself$_k$]$_i$ did Mary$_k$ lose t_i]?

Recall that the crucial aspect of reconstruction was its particular application at LF. That was problematic for various reasons; first, because it was evidence in support of the "level of representation" status of LF (as opposed to "component"; see Uriagereka forthcoming), and second, because it was an additional operation which implied the replacement of the trace. The copy theory provides an elegant alternative: since there is a copy, we do not need to put the moved element back to its base (first-Merge) position, all we have to care about is how much reconstruction we have. According to Chomsky (1993), economy factors dictate that only the restriction must reconstruct, the operator part of the moved element remaining upstairs, roughly as shown in (28):

(28) For [what x] Mary lost [x, picture of herself]

This said, some questions rapidly arise. As I see things, the most intriguing one concerns the nature of reconstruction itself: is it always possible? Chomsky (1993) restricts reconstruction to A'-Movement, which, at one level, makes sense: A-Movement creates no interpretive import, whereas the two positions created by A'-Movement instantiate an operator-variable relation.[30] In this sense, Chomsky (1993:205) adduces the following data indicating no reconstruction effects in the case of A-Movement:

[30] See Abels (2003), Barss (2001), Boeckx (2001, 2003a), Chomsky (2005b), Epstein & Seely (2006), Lasnik (1999), and Takahashi (1994) for related issues.

(29) a. $[_{CP}\ [_{TP}$ [The claim that John$_i$ was asleep]$_j$ seems to him$_i$ $[_{TP}\ t_j$ to be correct]]]
b. * I seem to him$_i$ to like John$_i$

The important thing to note about (29a) is the lack of Principle C effects, contrary to what one would be led to expect given our previous discussion. Chomsky's (1993) examples are well-taken, but they do not cover the whole range of facts: as Boeckx (1999) notes, binding and NPI licensing *qua* reconstruction reappear when the moved element is an indefinite:

(30) a. $[_{CP}\ [_{TP}$ [Pictures of himself$_i$]$_j$ seem to John$_i$ $[_{TP}\ t_j$ to be ugly]]]
b. $[_{CP}\ [_{TP}$ [Pictures of any artist]$_j$ seem to no critics $[_{TP}\ t_j$ to be revealing]]]

The tension between (29) and (30) is not easy to solve. Boeckx (1999) suggests that binding effects in (30) arise on-line, at the point where the experiencer is merged and the subject has not moved yet.[31]

The data, viewed in isolation, are puzzling, and different explanations have been offered in the recent literature. Lasnik (1999), for instance, suggests that A-Movement leaves no copy, whereas Fox (2000) argues that it leaves a trace (in the pre-Chomsky 1993 sense). As Boeckx (2001) argues, the mystery vanishes the minute one pays attention to what the difference between traces left by A and A' movements was in the GB-framework: the former do not have their Case checked, the latter do. Boeckx (2001) highlights that asymmetry in order to claim that copies left by A-Movement are "infected" (see Uriagereka 1998), and therefore unable to feed the interpretive component for reconstruction purposes.

It remains to assess the precise mechanism which decides what copy is going to be interpreted in, say, a long-distance extraction case. Consider (31), where the anaphor *himself* can take both *John* and *Peter* as its antecedent:

(31) Which picture of himself$_{\{i/j\}}$ did John$_i$ say that Peter$_j$ lost?

Assuming successive cyclic movement, the wh-phrase *which picture of himself* stops at SPEC-C of the embedded clause, as indicated in (32).[32]

(32) $[_{CP}$ [Which picture of himself]$_{-3}$ did John say $[_{CP}\ t_{-2}$ that Peter lost t_{-1}]]?

Unless stipulated, everything we have said so far predicts that any copy left by A'-movement can be interpreted at LF. Example (31) supports this prediction, as the available readings indicate. If the first (lowest) trace is interpreted, *Peter* binds the anaphor; if it is the second (intermediate) one, then *John* does.[33]

So far so good. Let us now take this excursus about binding as a theoretical background and consider how it deals with adjunct phenomena. As those familiar with the relevant literature know, we should start by considering (33), taken from Lebeaux (1991):

---

[31] Note that this does not explain why reconstruction effects –even if they arise on-line- only affect indefinites. Boeckx (2001) does offer an explanation for this, arguing that reconstruction only affects indefinites because only these undergo a process of literal LF reconstruction (i.e., a full-fledged lowering operation at LF).

[32] More accurately, it stops at every edge (see Chomsky 2000, 2001, 2005b), including every SPEC-*v**. I put aside the issue of whether movement always target SPEC-*v** and SPEC-C. See Rackowski & Richards (2005).

[33] I am sweeping some problematic cases under the rug, like Chomsky's (1993) preference principle which separates Principle-A (cliticization at LF) from Principle-B and Principle-C.

(33) a. $[_{CP}$ [Which claim [that John$_k$ made]]$_i$ did he$_k$ later deny t$_i$]?
b. * $[_{CP}$ [Which claim [that John$_k$ likes Mary]]$_i$ did he$_k$ deny t$_i$]?

In (33) there is a puzzling asymmetry: *John* does not trigger a Principle-C effect in (33a), but it does in (33b). Why? Many proposals have been invoked to solve this asymmetry, which is known as anti-reconstruction. While some authors have argued that depth of embedding is the key (see van Riemsdijk & Williams 1981), some others have argued that what counts is the semantic nature of the displaced element (see Heycock 1995). In parallel to these approaches to reconstruction puzzles, Lebeaux's (1991) account appealed to the argument vs. adjunct distinction, suggesting that since the latter class of elements (but not the former) does not need to appear (in s-selection terms) in base structures, it can be late-merged as the derivation goes along by means of generalized transformations.[34] Chomsky (2004) agrees with the basic facts, but notes the problems late-insertion analyses pose to cyclicity and the Extension Condition, so he resorts to the parallel plane idea: this allows him to preserve a strong form of cyclicity, while tackling the Principle-C effect.

> "Assume that like other operations, adjunction of α to β applies cyclically. β behaves throughout as if it were in a simple structure formed by set Merge […] What about Condition (C) at SEM? When X c-commands <α, β>, does it also c-command α and β? β was introduced by set Merge, and before α was adjoined to it, X c-commanded β. But the central property of adjunction of α to β does not change the properties of β. For β to lose some property when α adjoins to it would be a complication, an "imperfection". The relation c-command (X, β) is therefore not lost when α is adjoined to β: accordingly, X still c-commands β in <α, β>, as before adjunction. But extension of c-command to the adjoined element α would be a new operation, to be avoided unless empirically motivated. Happily, the empirical evidence disconfirms the complication." [Chomsky (2004:118)]

As Chomsky (2004) notes, Principle-C effects show up in sentences like (34a) and (34b), but this does not affect Lebeaux's (1991) main observation –anti-reconstruction effects only arise in adjuncts which have been fronted.

(34) a. * He$_k$ did not like the claim that John$_k$ made.
b. He$_k$ asked $[_{CP}$ [which claim that John$_{\{*k/j\}}$ made]$_i$ he$_j$ denied t$_i$]

(34b) is more interesting than (34a), for it is related to (35), also from Lebeaux (1991):

(35) a. $[_{CP}$ [Which paper that he$_j$ gave to Bresnan$_z$]$_i$ did every student$_j$ think that she$_z$ would like t$_i$]]?
b. * $[_{CP}$ [Which paper that he$_j$ gave to Bresnan$_z$]$_i$ did she$_z$ think that every student$_j$ would like t$_i$]?

In (35) we are dealing with both Principle C and variable binding. In (35a) the adjunct *that he gave to Bresnan*, contained within the wh-phrase headed by *which paper*, shows reconstruction effects in the SPEC-C position of *like*: from that position the pronoun *he* can

[34] See Stepanov (2001) for relevant discussion on late insertion of adjuncts.

be bound by *every student*, and still avoid the Principle C effect; in (35b), *every student* forces reconstruction in the base position, and this is fatal, due to Principle-C, not variable binding.

How does Chomsky's (2004) analysis account for the facts in (34) and (35)? He proposes a process of simplification (SIMPL) which is part of the Transfer operation applying at the phase level and restoring the normal structural relations (c-command, etc). Importantly, SIMPL is designed so that it can (but must not) dissociate PF and LF Transfer. That is, given any non-trivial chain, SIMPL can apply to one occurrence in the LF component and to another one in the PF component, thus explaining the facts in (35a), where Transfer effects are dissociated: in (35a), we spell-out (that is, Transfer to PF) the higher copy, whereas it is an intermediate one that gets shipped to LF for binding business.

> "There is an operation SIMPL that converts $<\alpha, \beta>$ to $\{\alpha, \beta\}$; in effect, it is part of Σ. Since SIMPL applies at the stage of the derivation at which Spell-Out S-O applies, it is also in effect part of S-O. We conclude, then, that it is part of the operation TRANSFER, which transfers the NS derivation (specifically, its last line) to both Φ and Σ. Suppose SIMPL is optional […] For overt movement […] optionality of SIMPL will have no effect at the PHON level because S-O does not apply to the trace in any event. But it might have an effect at SEM. Thus in such structures as [(34b)], application of SIMPL to the trace (copy) yields reconstruction effects, obviated if SIMPL applies only where it must: at the phase where S-O applies." [Chomsky (2004:118-119)]

This optional character of SIMPL also predicts the binding effect of sentences like (36), where it is the first (neither the last nor the intermediate one) copy which is simplified at LF.

(36) $[_{CP}$ [Which papers [that he$_k$ wrote]]$_i$ did every linguist$_k$ publish t_i]?

The facts fall into place as far as something like SIMPL is available. The asymmetry between variable binding and Principle-C effects is also captured due to the optional nature of this operation. There is, however, the logical doubt of whether the same results can be achieved without the operations of pair-Merge and SIMPL; or, in other words, whether we can question (and in fact deny) the ontological different status of these operations without empirical loss. Perhaps this is too premature an issue to pose, but since it is minimalist in its spirit, it is not out of place. In the next section I will offer more data showing reconstruction effects of adjuncts, but this time I will provide really relevant data, for note that Lebeaux's (1991) examples are no true adjuncts: they are adjuncts within arguments, which does not allow us to test the empirical adequacy of the parallel plane idea.

4. Two types of adjuncts

The goal of this section is to test the validity of Uriagereka's (2005) claim that modification can be formally expressed through pair-Merge (bona fide adjunction) and set-Merge (adjuncts merged as SPECs in a cascade structure), as indicated in (37), where α stands for the adjunct:

(37) a. True adjunction b. Specification

The facts to be considered can be classified depending on the degree of connectivity adjuncts display. At first glance, it seems that the system differentiates adjuncts on linear order grounds: adjuncts to the right show less connectivity than adjuncts to the left. It is important to note, before exploring the data in more detail, that this left vs. right asymmetry is predicted to vanish in the case of covert processes if Chomsky's (2004) SIMPL is correct; that is, if processes like NPI licensing or binding operate at the boundaries of LF (see Chomsky 2004, 2005b for discussion), it is somewhat pointless to make a distinction like (37): at LF the structure in (37a) would have been simplified, being formally identical to that of (37b). My own judgment, though, is that such a covert analysis is not forced upon us, for if NPI licensing and binding obtain under Agree, then those processes can take place within Narrow Syntax.

Obviously, the distinction in (37), as such, says nothing about sub-extraction from within adjuncts (Huang's 1982 Adjunct Condition). These dependents, analyzed either as SPECs or as true adjuncts, block all kind of sub-extraction.[35]

(38) *[$_{CP}$ [What books]$_i$ did John call Mary [$_{becauseP}$ because she had t_i]?

[35] As Juan Uriagereka (p.c.) notes, adjuncts of the sort studied by Browning (1987) allow sub-extraction:

(i) Who$_i$ did you go there [to visit t_i]

I hasten to add that I have no good explanation for (i) at this point, but it is not unlikely that some reanalysis process can be blamed. I suspect, to be precise about it, that the sequence 'go to visit' can be treated as a complex verbal predicate (i.e., a periphrasis).

There are similar cases of acceptable sub-extraction in Spanish, like (ii) and (iii) (presumably, parasitic gap structures). As can be seen, sub-extraction is possible if the verb within the adjunct clause is inflected in subjunctive mood or non-inflected.

(ii) *¿Qué libro$_i$ compraste t_i [$_{porP}$ porque Juan leyó t_i]? (Spanish)
What book bought-2SG for-that Juan read-3SG
'What book did you buy because Juan read?'

(iii) ¿Qué libro$_i$ compraste t_i [$_{paraP}$ para que Juan leyese t_i]? (Spanish)
What book bought-2SG to that Juan read-SUBJ-3SG
'What book did you buy for Juan to read?'

(iv) ¿Qué libro$_i$ compraste t_i [$_{paraP}$ para leer t_i]? (Spanish)
What book bought-2SG to to-read
'What book did you buy to read?'

The Dana Scully character of the X-Files series has provided me with the next intriguing datum, for which I have no explanation either:

(v) La verdad$_i$ [$_{CP}$ [que$_i$] tanto hemos luchado [$_{porP}$ por descubrir t_i]] (Spanish)
The truth that so-much have-1PL fought to to-discover
'The truth we have fought so much to discover'

Curiously, if the quantifier *tanto* (Eng. *so much*) is dropped extraction is worse, perhaps indicating that the structures share some properties in common with result clauses.

(vi) ??La verdad$_i$ [$_{CP}$ [que$_i$] hemos luchado [$_{porP}$ por descubrir t_i]] (Spanish)
The truth that have-2PL fought for to-discover
'The truth that we have fought to discover'

All other things being equal, an explanation for the Adjunct Condition should lie in the specifics of locality, about which there is no consensus. Reducing all proposals I am familiar with to two is a non-trivial cut, but I think it is fair to argue that there are two main views: sub-extraction is said to be impossible due either to phrase structure or to agreement facts. Let me elaborate: while some scholars argue that structural constraints rule sub-extraction (see Boeckx 2003a; Chomsky 2004; and Uriagereka 2003 for the parallel plane idea; Chomsky 2005b about locality problems with phase edges; and Uriagereka 1999 on the notion of command unit and Multiple Spell-out), others believe that the mechanics of Agree systems draw the line (see Boeckx 2003a; and Rackowski & Richards 2005). The possibility that something like Chomsky's Activity Condition is relevant here strikes me as an appealing one; in particular, it would be interesting if (39), or some variant, was the relevant principle.[36]

(39) **Activity Condition on Sub-extraction**
Only active elements (those that have not been assigned Case) are transparent to sub-extraction operations.

In the case of embedded CPs, which are not usually strong islands, there seems to be empirical evidence that Agree processes are complex (or "multiple", in the sense of Hiraiwa 2001, 2005), since sub-extraction of wh-phrases requires for the Probe to previously Agree with the whole CP (see Boeckx 2003a; Rackowski & Richards 2005; and Uriagereka 2002). If we apply the same logic to adjuncts, the prediction is clear: due to the same reason that adjuncts project no label, adjuncts cannot be Goals for φ-Probes. The question that immediately arises is whether adjuncts can establish a (long-distance) Probe-Goal dependency which does not need to invoke a Match-Agree procedure? This question, though, cannot be formulated if adjuncts are in a parallel plane, for then they do not enter into c-command relations; if, however, they are not (being a SPEC), we can formulate this question, and the Spanish data in (40) become extremely relevant in this respect.

(40) a. ¿[Con cuál de sus$_z$ novias]$_i$ crees que va más al cine
with which of his girlfriends think-2SG that go-3SG more to-thecinema
todo guaperas$_z$ t$_i$?
every handsome-boy
'Which of his girlfriends do you think every good-looking-guy goes to the cinema with the most?'

b. ¿[Para cuántos de sus$_z$ familiares]$_i$ ha comprado regalos cada
for how-many of his relatives have-3SG bought presents every
chico$_z$ t$_i$?
boy
'For how many of his relatives has every boy bought presents?'

[36] Again, this poses many questions which I cannot address due to space restrictions. See Gallego (2006, in progress) for more detailed discussion. Everything boils down to what being a "barrier" (or "opaque domain") is. Obviously, being a SPEC alone will not do if the facts noted by Uriagereka (1988) and Chomsky (2005b) are correct (some SPECs allow sub-extraction). It remains to be discovered, then, whether the parallel plane and the Activity Condiction accounts are independent from each other or are finally reducible to one.

c. ¿[Por cuál de sus$_z$ crímenes]$_i$ ha sido encarcelado todo
for which of his crimes have-3SG been put-into-jail every
mafioso$_z$ t_i?
gangster
'Which of his crimes do you think every gangster has been put into jail for?'

The examples in (40) show that true wh-adjuncts (those, which unlike Lebeaux's 1991, are not within arguments) can reconstruct in a position lower than the external argument[37], a conclusion incompatible with Chomsky's (2004) analysis: being pair-Merged, adjuncts must be out-of-sight for any displacement operation.[38, 39]

Facts concerning scope are not so problematic. In principle, the analyses in (37) allow for scopal and non-scopal readings to obtain, but, as noted by Ernst (2000, 2002), sentences like those in (41) do not differ on truth-conditional grounds. Thus, in (41), the adjuncts *in the kitchen* and *hungrily* do not seem to scope over the VP-chunk they adjoin to.[40, 41]

(41) a. Carol ate the fish hungrily in the kitchen.
b. Carol ate the fish in the kitchen hungrily.

The sentences in (41) get what I will call, in the spirit of Uriagereka (2003) and Martin & Uriagereka (2000), a Markovian reading, perfectly fitting with neo-Davidsonian treatments of adjuncts as predicates of the event (see Davidson 1967; Herburger 2000; and Parsons 1990):

(42) [∃ e: eat (e) & Agent (Carol,e) & Theme (the fish,e) & hungrily (e) & in-the-kitchen (e)]

Notice that (42) can account for the entailment patterns of (43): (43a) entails both (43b) and (43c), which, together, they do not entail (43a) (see Parsons 1990).

---

[37] This constitutes a problem for accounts in which wh-adjuncts are base-generated in SPEC-C. See Boeckx (2003a:91-97) for relevant discussion.

[38] I owe this observation to Cedric Boeckx (p.c.).

[39] Noam Chomsky (p.c.) notes that the wh-constituents in (40) are not adjuncts, and can be analyzed as SPECs in a cascade structure à la Pesetsky (1995). This conclusion is identical and compatible to Uriagereka's (2005), confirming the necessity for the duality of (37).

[40] Things are different in Cinque's (1999) proposal, which, due to theory internal reasons, analyzes all VP adjuncts as SPECs in *v**P shells with massive "snowballing" movements. Cinque (1999:28) does provide, however, cases in which adjuncts display strong scope effects; note, crucially, that his examples contain quantified adjuncts, which surely blur the picture.

(i) He attended classes every day of the week in a different university.
(*every day of the week* scopes over *in a different university*)
(ii) He attended classes in each university on a different day of the week.
(*in each university* scopes over *on a different day of the week*)

[41] In this respect, a reviewer is concerned about sequences like (i) and (ii), both of which contain adjuncts (relative clauses) displaying scope (embedding) properties:

(i) * [The man [the woman [the child knows] loves] comes]
(ii) [The man [the woman knows] comes]

I think neither (i) nor (ii) pose any problem for the scopeless nature of adjuncts, since, as we are about to see, adjuncts can also manifest scope effects. As for the particular contrast between (i) and (ii), I have nothing especially deep to say, other than it might follow from memory limitations –plausibly, when the pattern 'The N that John V_1 V_2' is reached, computational restrictions would force the restatement of expressions in paratactic terms, as Chomsky (2005b, 2006) suggests.

(43) a. [∃ e: eat(e) & Agent(Carol,e) & Theme(the fish,e) & hungrily(e) & in-the-kitchen(e)]
b. [∃ e: eat(e) & Agent(Carol,e) & Theme(the fish,e) & hungrily(e)]
c. [∃ e: eat(e) & Agent(Carol,e) & Theme(the fish,e) & in-the-kitchen(e)]

All this does not mean that scope (non-Markovian) readings are impossible. (44), taken from Ernst (2000:353), receives precisely such a reading.

(44) a. They run fast awkwardly, but run slowly smoothly.
b. They play soft well enough, but play loudly pretty poorly.

As Irurtzun & Gallego (2006) note, the semantics we want for this type of expressions is roughly that of conditionals. In our paper, we propose the semantics of (45), which should be read as follows: 'all events of they running fast are events of they running fast awkwardly'.

(45) [∀ e: run(e) & Agent(they,e) & fast(e)] awkwardly(e)

In Irurtzun & Gallego (2006) it is claimed that pair-Merge can be used to provide this non-Markovian reading too, but this is unlikely: it would not explain why the same operation can provide two different readings –there must be something formally different. Here I would like to propose that non-Markovian readings arise when adjuncts are merged as SPECs. To keep things simple, I will use the sentence in (46), which can receive both interpretations, as indicated in (47):

(46) John works in Barcelona.

(47) a. [∃ e: work(e) & Agent(John,e)] in-Barcelona(e) *Markovian Reading*
b. [∀ e: work(e) & Agent(John,e)] in-Barcelona(e) *Non-Markovian Reading*

Syntactically, I would like to argue that the non-Markovian reading arises if the adjunct *in Barcelona* is an outer-SPEC-*v**, as indicated in (48).

(48) [$_{v*P}$ [in Barcelona] [$_{v*P}$ John [$_{v*'}$ *v** [$_{VP}$ work]]]]

How is *in Barcelona* merged as a SPEC? [42] I will assess this point by the end of this section in a more detailed way, suffice it to say at this point that this merger is triggered by an EPP feature that *v** is endowed with (an edge feature, in Chomsky's 2005b terms). I consider (48) as a structure with a semantic import on the outcome, similar, mutatis mutandis, to the case of Object Shift explored by Chomsky (2001) and to the analysis of focus developed by Irurtzun (2005).[43] Accordingly, my analysis of non-Markovian adjuncts presupposes a quantificational treatment which is consistent with the semantics advanced by Irurtzun & Gallego (2006).

[42] At first glance, there are two possible strategies: either *in Barcelona* is first-Merged as a SPEC or else it moves there from a true adjunction position? The latter possibility is barred if Probe-Goal dependencies need c-command to go through, as Cedric Boeckx (p.c.) makes me note, so I will assume the former. See Gallego (in progress) for discussion.

[43] See Belletti (2004) and Rizzi (1997) for related ideas in a different framework.

What about NPI licensing and binding? Consider NPI first. As noted by Uriagereka (2003), examples like (49) reinforce the hypothesis we started this section with: adjuncts to the right behave as true adjuncts; adjuncts to the left behave as SPECs:

(49) a. * [$_{CP}$ A priest$_i$ can <[$_{v*P}$ t$_i$ tell *any* secret of confession] [under no circumstances]>]
b. [$_{CP}$ [Under no circumstances] can a priest tell *any* secret of confession]

In (49), only when placed to the left, the negative adjunct *under no circumstances* can c-command (and license) the NPI *any secret of confession*. What happens if it is the adjunct that contains an NPI (see Larson 2004; Marantz 2003: and Pesetsky 1995)? Surprisingly enough, we find the opposite pattern:

(50) a. María no habló con Juan en ninguna biblioteca. (Spanish)
María not talked-3SG with Juan in any library
'María did not talk to Juan in any library'
b.??En ninguna biblioteca, María no habló con Juan. (Spanish)
In any library, María not talked-3SG with Juan
'In any library, María did not talk to Juan'

(50b) is expected: since no negative element c-commands the adjunct, the NPI is not licensed.[44] (50a) is unexpected, but assuming NPI licensing to operate at LF, a SIMPL based analysis could do. The same could be said about binding within adjuncts (see Pesetsky 1995):

(51) a. He [$_{v*P}$ <[cleaned every room$_i$] [for its$_i$ owner]>]
b. She [$_{v*P}$ <[gave a present to the children$_i$] [on each other$_i$'s birthdays]>]
c. A treat was [$_{v*P}$ <[given to every pet$_i$] [by its$_i$ owner]>]

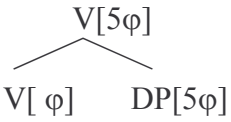
The binding effects in (51) can be analyzed in different ways: following work by Howard Lasnik, Chomsky (2005b) argues that objects always raise to a case checking position he identifies as SPEC-V from which they c-command into the adjunct after Transfer (when SIMPL has applied).[45] But do the facts in (51) necessarily follow from binding taking place at LF, as Chomsky (2005b) suggests, or can they be accommodated by treating adjuncts as SPECs? Note that we are facing the same paradox again and again: what decides whether modifiers are merged as true adjuncts (simplified at LF) or as SPECs? Here I would like to relate this issue to Boeckx's (2002) labelling proposal. As the reader may recall, the gist of Boeckx's (2002) analysis is that labels correspond to a feature shared by the SOs undergoing Merge. For the purposes of this paper I will assume that shared features reduce to two: EPP and φ-features.[46] Things being so, the merger of a verb and its object (Probe and Goal respectively) would be possible due to the φ-features they share, roughly as in (52).[47]

[44] The expression is not fully out for some speakers. I put aside the relation between (50b)'s degraded status and the topicalization of non-specific indefinite DPs.

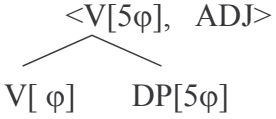
[45] I put aside the details of Larson's (2004) analysis, which, in order to capture the NPI, binding, linear order and focus properties of adjuncts, assumes that adjuncts undergo Merge with the verb before arguments do.

[46] This does not exhaust Boeckx's (2002) own typology, which includes Φ, Θ, and Q features. I discuss those options in Gallego (2007).

[47] I take it, much in line with what is generally assumed ever since the Split INFL Hypothesis, that there are functional nodes containing bundles of φ-features belonging to subject and object. With Chomsky (2005b), I assume that C and T contain subject φ-features, whereas *v** and V do object φ-features.

(52)


In (52) I assume that the φ-features of the verb are valued under Match-and-Agree with those of the object, and that the label (i.e., feature) which projects is that of the verb because it is the one that needs valuation –it is the Probe. In this way, we ensure that the resulting syntactic object has a verbal nature.[48] The same logic naturally extends to the merger of *v** and the external argument (which, as suggested by Chomsky 2005b, might give rise to an unlabeled structure). What about adjuncts? It is plausible enough that they do not share features with any SO, but if so, how must they be introduced? The duality of (37) can be translated by saying that adjuncts are merged by projecting no label at all (pair-Merge) or by projecting an EPP-induced label. The first possibility gives us a Match without Agree, in Boeckx's (2002:29) system: "just like we find Match without Agree in cases of non-trivial chains (displacement), so we expect to find Match without Agree under basic concatenation operations". I assume Boeckx's (2002) proposal except for the Matching part –no feature can be matched in adjunction. Things being so, adjuncts would be introduced as in (53):

(53)


(53) seems to be able to explain the low level of connectivity of adjuncts to-the-right, but more is needed: we have seen empirical evidence suggesting that, whatever the ultimate reason, adjuncts can be merged as SPECs. I want to argue that in these cases, the label is EPP-driven,[49] as in merger operations with phase heads *v** and C in order to create outer-SPECs. In truth, there is no feature sharing in these cases either (unless we stipulate them, as in the cartographic project of Rizzi 1997 and his associates): the phase head is just endowed with an indiscriminate EPP feature acting as a Probe, and no feature of the Goal is matched (in plain English: it makes no sense to postulate a valued counterpart for the EPP) so I assume that one of the Probe's features directly projects. The key question: which one? Here I will assume φ-features percolate in this case as well.[50]

---

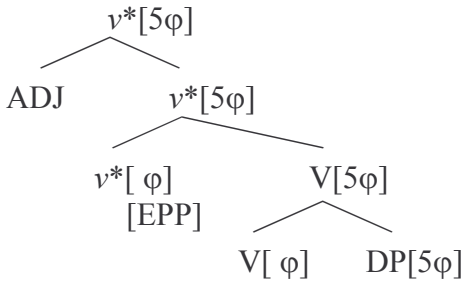
[48] This endorses Chomsky's (2001) and Boeckx's (2002) idea that categorial features can be dispensed with (see Marantz 1997). The fact that the features which project are uninterpretable in the verb need not be a problem (see Chomsky 1995, 2000 for discussion about the projection of labels of uninterpretable φ-features) as long as we stick to the leading role of valuation (and not interpretability) within computation. See Pesetsky & Torrego (2004).

[49] This is actually Chomsky's (2005b) claim, but while he generalizes an EPP-driven Merge, I restrict it to the creation of SPECs (in the sense of Chomsky 2000, 2001):

> "For an LI to be able to enter into a computation, merging with some SO (and automatically satisfying SMT), it must have some property permitting this operation. A property of an LI is called a feature, so an LI has a feature that permits it to be merged. Call this the edge-feature (EF) of the LI. If an LI lacks EF, it can only be a full expression in itself; an interjection".
>
> [Chomsky (2005b:6)]

[50] In (54) I put aside the merger of the external argument.

(54)



Notice that, in (54), we are forced to assume that some property (some feature) of *v** projects: that is the only way of capturing the verbal nature of the entire syntactic object, as well as the fact that the ADJ is merged as a SPEC, occupying *v**'s edge. The process is similar in its spirit to Chomsky's (2001) analysis of Object Shift and the idea that semantic effects arise at phase edges, since in these cases the resulting configuration has semantic consequences too: binding, non-Markovian readings, and NPI licensing are possible only under (54). Granted, some cases of binding (51) and NPI licensing (50) go through without (54) if Agree processes can operate at LF (see Nissenbaum 2000 and Pesetsky 2000), but if Agree only operates during computation, then we do need (54) in all such cases.

Finally, something must be said about linear order. Martin & Uriagereka (2000) and Uriagereka (2003) discuss this point at length, noting the tension that Kayne's (1994) LCA creates when adjunction comes to the fore. They conclude that there are three possible solutions: a) LCA is wrong; b) LCA is not wrong, and there are massive "snowballing" movements (see Cinque 1999); c) no linear order can be decided by standard methods. Martin & Uriagereka (2000) endorse (c), building on the fact that adjuncts, when pair-Merged, do not enter into c-command relations, and Kayne's (1994) LCA becomes useless.[51]

By exploiting the notion of Numeration (see Chomsky 1995, 2000, 2001), Uriagereka (2003) suggests a dynamic activation procedure that derives linear order and scope effects. In this paper, I will assume Uriagereka's (2003) proposal, whose details are as follows:

(55) **Syntactic Activation**
A syntactic object SO is activated when it leaves a Numeration NUM and enters a derivational workspace D_{WS}.

(56) **Consequences of derivational activation**
A modifier's derivational activation directly determines its linear order and scope.

(55) and (56) are not principles, but rather particular assumptions about the nature of derivational dynamics. Together, (55) and (56) imply that if adjunct A is "activated" (that is, if adjunct A leaves the NUM, entering the derivation proper) before adjunct B, adjunct A precedes adjunct B, no more no less. This assumption does not have to be made in the case of arguments, for these can piggy-back on c-command relations created during the computation. In the case of bona fide adjuncts c-command is unavailable, so the system resorts to an internal device (activation timing) in order to yield the desired output.

[51] The problem could go away if SIMPL works as Chomsky (2004) argues. If c-command can be restored at the Transfer point (due to SIMPL), then PF and LF could piggy back on that information to create both scope effects and linear order. See Gallego (2006) for a dynamic account of linear order for pair-Merged adjuncts different from Uriagereka's (2003).

Adjuncts which are formally SPECs, on the other hand, do not need to run the same fate – Kayne's (1994) LCA is enough. A red flag must be raised, nevertheless: even though linear order can be decided at PF, we are forced to assume that all arguments must abandon their first-Merge position (escaping from the VP; see Alexiadou & Anagnostopoulou 2001) before Transfer –that is the only way for adjuncts in SPEC-*v** to appear to-the-right.[52]

5. Conclusions

In the preceding lines I have explored some aspects of Bare Phrase Structure (BPS) from a minimalist point of view. Following Uriagereka (2005, forthcoming) I have assumed a connection between BPS and the Chomsky Hierarchy, paying special attention to the type 3 of grammars, which I have taken to embody Chomsky's (2004) pair-Merge. From that onwards, I have focused on one of the most exciting and obscure aspects of current syntactic theory: adjuncts. It has been my intention to derive all the properties of these syntactic dependents from a notion which I consider as the backbone of contemporary syntactic theory: phrase structure. As we have seen, there is robust evidence suggesting a non-uniform treatment for adjuncts, as argued by Uriagereka (2005, forthcoming). This duality seems to me to be unavoidable, and suggests that Cinque's (1999) analysis cannot be the whole history, as correctly noted by Boeckx (2003b):

> "It is fair to say that adjuncts are not specifiers; they don't behave on a par: Adjuncts iterate, specifiers, don't; specifiers overwhelmingly surface on the left, many adjuncts on the right (in SVO languages); some specifiers allow for extractions, adjuncts never do; specifiers license anaphors, adjuncts never do. The list could go on and on. Specifiers and adjuncts just are two different species. Collapsing them structurally loses any hope to distinguish them."
>
> [Boeckx (2003b:98-99)]

It remains to be found out why things are this way. That is: it remains to "understand" why this duality (if correct) arises. I have sketched a solution that is consistent with the interface-driven nature of adjuncts (see Boeckx 2003b; Ernst 2002; Chomsky 2004): adjuncts are SPECs depending on an effect on the outcome.[53] Perhaps this is a premature (or even wrong) conclusion, but the hypothesis appears to be compatible with almost everything we have seen: adjuncts usually happen to be SPECs whenever some interface-driven process is at stake (e.g., feature valuation in cases of NPI and binding, and also quantificational structures when non-Markovian readings arise). The connection may be, to repeat, too hasty, but it is worth exploring inasmuch as it satisfies the leading intuition that syntax is an optimal solution to interface requirements.

[52] Obviously, nothing extra needs to be said for wh-moved or topicalized adjuncts –these are also SPECs.

[53] As Cedric Boeckx (p.c.) correctly points out, this formal ambivalence (pair-Merge vs. set-Merge as a SPEC in cascade) resembles the optional labeling mechanism for adjuncts discussed by Hornstein et al. (2005): although adjunction does not normally yield labels, these can be created for computational operations with an effect on the outcome (e.g., VP topicalization).

Acknowledgements

For comments, helpful discussion, and suggestions I would like to thank Sergio Balari, Cedric Boeckx, Noam Chomsky, Norbert Hornstein, Wolfram Hinzen, Aritz Irurtzun, and Andrea Moro. Special thanks go to Juan Uriagereka, whose work inspires mine.

Ángel J. Gallego
Universitat Autònoma de Barcelona
angel.gallego@uab.es
http://seneca.uab.es/ggt/membres/gallego.htm

References

Abels, K. (2003). *Successive cyclicity, anti-locality, and adposition stranding*. Ph.D. Dissertation, U.Conn.
Alexiadou, A. & E. Anagnostopoulou (2001). The subject *in situ* generalization, and the role of case in driving computations. *Linguistic Inquiry* 32, pp. 193-231.
Barss, A. (2001). Syntactic reconstruction effects. M. Baltin & C. Collins (eds.), *The handbook of contemporary syntactic theory*. Blackwell, Malden (Mass.), pp. 670-696.
Belletti, A. (2004). Aspects of the low IP area. L. Rizzi (ed.), *The structure of CP and IP. The cartography of syntactic structures (vol. 2)*. Oxford University Press, Oxford, pp. 16-51.
Boeckx, C. (1999). Conflicting c-command requirements. *Studia Linguistica* 53, pp. 227-250.
Boeckx, C. (2001). Scope reconstruction and A-Movement. *Natural Language and Linguistic Theory* 19, pp. 503-548.
Boeckx, C. (2002). Elemental relations in syntax. Ms., UMD.
Boeckx, C. (2003a). *Islands and chains. Resumption as stranding*. John Benjamins, Amsterdam.
Boeckx, C. (2003b). Review to *The syntax of adjuncts*. *Studies in the Linguistic Sciences* 32, pp. 95-101.
Browning, M. (1987). *Null operator constructions*. Ph.D. Dissertation, MIT.
Chametzky, R. (2000). *Phrase structure: from GB to minimalism*. Blackwell, Malden (Mass.).
Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory* 2, pp. 113-124.
Chomsky, N. (1963). Formal properties of grammar. , R.D. Luce, R.R. Bush & E. Galanter (eds), *Handbook of mathematical psychology (vol.2)*, Wiley, New York, pp. 323-418.
Chomsky, N. (1970). Remarks on nominalization. R.A. Jacobs & P.S. Rosembaum (eds.), *Readings in English transformational grammar*. Georgetown University Press, Washington, pp. 184-221.
Chomsky, N. (1981). *Lectures on government and binding*. Foris, Dordrecht.
Chomsky, N. (1986). *Barriers*. MIT Press, Cambridge (Mass.).
Chomsky, N. (1993). A minimalist program for linguistic theory. K. Hale & S. J. Keyser (eds.), *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*. MIT Press, Cambridge (Mass.), pp. 1-52.
Chomsky, N. (1995). *The minimalist program*, MIT Press, Cambridge (Mass.).
Chomsky, N. (2000). Minimalist inquiries: the framework. R. Martin, D. Michaels, & J. Uriagereka, *Step by step. Essays on minimalist syntax in honor of Howard Lasnik*. MIT Press, Cambridge (Mass.), pp. 89-155.
Chomsky, N. (2001). Derivation by phase. M. Kenstowicz (ed.), *Ken Hale: a life in language*. MIT Press, Cambridge (Mass.), pp. 1-52.
Chomsky, N. (2004). Beyond explanatory adequacy. A. Belletti (ed.), *Structures and beyond. The cartography of syntactic structures (vol. 3)*. Oxford University Press, Oxford, pp. 104-131.
Chomsky, N. (2005a). Three factors in language design. *Linguistic Inquiry* 36, pp. 1-22.
Chomksy, N. (2005b). On phases. Ms., MIT.
Chomsky, N. (2006). Approaching UG from below. Ms., MIT.
Chomsky, N. & H. Lasnik (1995). The theory of principles and parameters. Chomsky, N., *The minimalist program*. MIT Press, Cambridge (Mass.), pp. 13-127.
Cinque, G. (1999). *Adverbs and functional heads. A cross-linguistic perspectivy*. Oxford University Press, New York.

Collins, C. (2002). Eliminating labels. S. Epstein & T. Seely (eds.), *Derivation and explanation in the minimalist program*. Blackwell, Malden (Mass.), pp. 106-132.
Davidson, D. (1967). The logical form of action sentences. *Essays on actions and events*. Claredon Press, Oxford, pp. 105-122.
Donati, C. (2006). On wh-head movement. L. Cheng & N. Corver (eds.), *Wh-movement: moving on*. MIT Press, Cambridge Mass., pp. 21-46.
Epstein, S. & D. Seely (2006). *Derivations in minimalism*. Cambridge University Press, Cambridge.
Ernst, T. (2000). Manners and events. C. Tenny & J. Pustejovsky (eds.), *Events as grammatical objects*. CSLI Publications, Stanford, pp. 335-358.
Ernst, T. (2002). *The syntax of adjuncts*. Cambridge University Press, Cambridge.
Frampton, J. & S. Gutmann (2000). Agreement is feature sharing. Ms., Northeastern University.
Fukui, N. (2001). Phrase structure. M. Baltin & C. Collins (eds.), *The handbook of contemporary syntactic theory*, Blackwell, Malden (Mass.), pp. 374-406.
Fox, D. (2000). *Economy and semantic interpretation*. MIT Press, Cambridge (Mass.).
Gallego, A.J. (2006). Dynamic phrase structure. Ms., UAB.
Gallego, A.J. (2007). *Phase theory and parametric variation*. Ph.D. Dissertation, UAB.
Hale, K. & S.J. Keyser (2002). *Prolegomenon to a theory of argument structure*. MIT Press, Cambridge (Mass.).
Harley, H. (2003). How do verbs get their name? Ms., University of Arizona.
Heycock, C. (1995). Asymmetries in reconstruction. *Linguistic Inquiry* 26, pp. 547-570.
Herburger, E. (2000). *What counts: focus and quantification*. MIT Press, Cambridge (Mass.).
Hiraiwa, K. (2001). Multiple agree and the defective intervention constraint in Japanese. O. Matushansky., A. Costa, J. Martín-González, L. Nathan, & A. Szczegielniak (eds.), *MIT working papers in linguistics* 40, pp. 67-80.
Hiraiwa, K. (2005). *Dimensions of symmetry in syntax: agreement and clausal architecture*. Ph.D. Dissertation, MIT.
Hornstein, N. (2005). What do labels do? Some thoughts on the endocentric roots of recursion and movement. Ms., UMD.
Hornstein, N., J. Nunes, & P. Pietroski (2005). Some thoughts on adjunction. Ms., UMD/USP.
Hornstein, N. & J. Uriagereka (2002). Reprojections. S. Epstein & D. Seely (eds.), *Derivation and explanation in the minimalist program*. Blackwell, London, pp. 107-132.
Huang, J. (1982). *Logical relations in Chinese and the theory of grammar*. Ph.D. Dissertation, MIT.
Irurtzun, A. (2005). Focus and clause structuration in the minimalist program. Boeckx, C. (ed.), *Minimalist Essays*. John Benjamins, Amsterdam, 68-96.
Irurtzun, A. & A.J. Gallego (2006). Consequences of pair-merge at the interfaces. Ms., UAB – UPV/EHU & IKER/CNRS.
Kayne, R. (1994). *The antisymmetry of syntax*. MIT Press, Cambridge (Mass.).
Larson, R. (2004). Sentence-final adverbs and 'scope'. M. Wolf & K. Moulton (eds.), *Proceedings of NELS 34*. U.Mass., GLSA, pp. 23-43.
Lasnik, H. (1999). Chains of arguments. S. Epstein & N. Hornstein (eds.), *Working minimalism*. MIT Press, Cambridge (Mass.), pp. 189-215.
Lebeaux, D. (1991). Relative clauses, licensing, and the nature of the derivation. S. Rothstein (ed.), *Perspectives on phrase structure: heads and licensing*. Academic Press, San Diego, pp. 209-239.
Marantz, A. (1984). *On the nature of grammatical relations*. MIT Press, Cambridge (Mass.).
Marantz, A. (1997). No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. A. Dimitriadis & L. Siegel (eds.), *U. Penn working papers in linguistics* 4, pp. 201-225.
Marantz, A. (2003). Objects out of the lexicon: argument-structure in the syntax. Ms., MIT.
Martin, R. & J. Uriagereka (2000). Collapsed waves in natural language. Ms., UT & UMD.
Mateu, J. (2002). *Argument structure: relational construal at the syntax-semantic interface*. Ph.D. Dissertation, UAB.
May, R. (1985). *Logical form. Its structure and derivation*. MIT Press, Cambridge (Mass.).
Moro, A. (2000). *Dynamic antisymmetry*. MIT Press, Cambridge (Mass.).
Nissenbaum, J. (2000). *Investigations of covert phrase movement*. Ph.D. Dissertation, MIT.
Nunes, J. (2004). *Linearization of chains and sideward movement*. MIT Press, Cambridge (Mass.).
Parsons, T. (1990). *Events in the semantics of English*. MIT Press, Cambridge (Mass.).
Pesetsky, D. (1995). *Zero syntax. Experiencers and cascades*. MIT Press, Cambridge (Mass.)..
Pesetsky, D. (2000). *Phrasal movement and its kin*, MIT Press, Cambridge (Mass.).

Pesetsky, D. & E. Torrego (2004). The syntax of valuation and the interpretability of features. Ms., MIT/U.Mass. Boston.
Rackowski, A. & N. Richards (2005). Phase edge and extraction: a Tagalog case study. *Linguistic Inquiry* 4, pp. 565-599.
Richards, N. (2001). A distinctness condition on linearization. Ms., MIT.
Riemsdijk, H. van & E. Williams (1981). NP-structure. *Linguistic Review* 1, pp. 171-217.
Rizzi, L. (1990). *Relativized minimality*. Cambridge (Mass.), MIT Press.
Rizzi, L. (1997). The fine structure of the left periphery. Haegeman, L. (ed.), *Elements of grammar. Handbook in generative syntax*. Kluwer, Dordrecht, pp. 281-337.
Starke, M. (2004). On the inexistence of specifiers and the natures of heads. A. Belletti (ed.), *Structures and Beyond. The Cartography of Syntactic Structures (vol. 3)*. Oxford, Oxford University Press, pp. 252-268.
Stepanov, A. (2001). *Cyclic domains in syntactic theory*. Ph.D. Dissertation, U.Conn.
Takahashi, D. (1994). *Minimality on movement*. Ph.D. Dissertation, U.Conn.
Tenny, C. (1994). *Aspectual roles and the syntax-semantics interface*. Kluwer, Dordrecht.
Uriagereka, J. (1988). *On Government*. Ph.D. Dissertation, U.Conn.
Uriagereka, J. (1998). *Rhyme and reason. An introduction to minimalist syntax*. MIT Press, Cambridge (Mass.).
Uriagereka, J. (1999). Multiple spell-out. S. Epstein. & N. Hornstein (eds.), *Working minimalism*. MIT Press, Cambridge (Mass.), pp. 251-282.
Uriagereka, J. (2002). Minimal restrictions on Basque movements. J. Uriagereka, *Derivations: Exploring the Dynamics of Syntax*. Routledge, London/New York,pp. 86-114.
Uriagereka, J. (2003). Pure adjuncts. Ms., UMD.
Uriagereka, J. (2004). Consequences of multiple spell-out. Ms., UMD.
Uriagereka, J. (2005). A Markovian syntax for adjuncts. Ms., UMD.
Uriagereka, J. (forthcoming). *Syntactic anchors*. Cambridge, Cambridge University Press.