MIT Press, Cambridge, Ma.
Chung, Sandra and James McCloskey (1987). Government, barriers, and small clauses in Modern Irish. *Linguistic Inquiry* 18.2: 173–237.
Chung, Sandra (1990). VP's and verb movement in Chamorro, *Natural Language and Linguistic Theory* 8: 559–619.
Craig, Colette (1977). *The Structure of Jacaltec.* University of Texas Press, Austin, Tx.
Doherty, Cathal (1996). Clausal structure and the Modern Irish copula, in *Natural Language and Linguistic Theory* 14: 1–46.
Doherty, Cathal (1997). Predicate initial constructions in Irish, in *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*, ed. by Brian Agbayani & Sze-Wing Tang, 81–95. CSLI, Stanford, Ca.
Duffield, Nigel (1995). *Particles and Projections in Irish Syntax.* Kluwer, Dordrecht.
Emonds, Joseph (1976). *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations.* Academic Press, New York.
Gazdar, Gerald, Geoffrey Pullum, and Ivan Sag (1982). Auxiliaries and related phenomena in a restrictive theory of grammar, *Language* 58: 591–638.
Halle, Morris, and Alec Marantz (1993). Distributed morphology and the pieces of inflection, in *The View from Building 20*, ed. by Kenneth Hale and Samuel Jay Keyser, 11–76. MIT Press, Cambridge, Ma.
Halpern, Aaron (1995). *On the Placement and Morphology of Clitics.* CSLI, Stanford, Ca. (1992 Stanford PhD dissertation).
Hinterhölzl, Roland (1997). A VO-based approach to verb raising, in *Proceedings of the North East Linguistic Society* 27, ed. by Kiyomi Kusumoto, 187–202. GLSA, University of Massachusetts, Amherst, Ma.
Hoekstra, Eric (1991). *Licensing Conditions on Phrase Structure,* PhD dissertation, Rijksuniversiteit Groningen, Grodil, Groningen.
Holmberg, Anders (1997). The true nature of Holmberg's Generalization, in *Proceedings of the North East Linguistic Society* 27, ed. by Kiyomi Kusumoto, 203–217. GLSA, University of Massachusetts, Amherst, Ma.
Kaye, Alan (1987). Arabic, in *The World's Major Languages*, ed. by Bernard Comrie, 664–685. Oxford University Press, Oxford.
Kayne, Richard (1994). *The Antisymmetry of Syntax.* MIT Press, Cambridge, Ma.
Kayne, Richard (1998). Overt vs. covert movement, to appear in *Syntax* 1.2.
Koopman, Hilda (1994). Licensing heads, in *Verb Movement*, ed. by David Lightfoot and Norbert Hornstein, 261–296. Cambridge University Press, Cambridge.
Kornai, Andras, and Geoffrey Pullum (1990). The X-bar theory of phrase structure, *Language* 66: 24–50.
Massam, Diane, and Carolyn Smallwood (1997). Essential features of predication in English and Niuean, in *Proceedings of the North East Linguistic Society* 27, ed. by Kiyomi Kusumoto, 263–272. GLSA, University of Massachusetts, Amherst, Ma.
McCloskey, James (1996). On the scope of verb movement in Irish, *Natural Language & Linguistic Theory* 14: 47–104.
Miller, Philip (1991). *Clitics and Constituents in Phrase Structure Grammar.* PhD dissertation, Rijksuniversiteit Utrecht.
Mosel, Ulrike, and Even Hovdhaugen (1992). *Samoan Reference Grammar.* Scandinavian University Press, Oslo.
Muysken, Pieter (1982). Parameterizing the notion 'head', *Journal of Linguistic Research* 2: 57–75.
Pollard, Carl and Ivan Sag (1994). *Head-Driven Phrase Structure Grammar.* The University of Chicago Press, Chicago, Ill.
Pulleyblank, Douglas and Akinbiyi Akinlabi (1988). Phrasal morphology in Yoruba, *Lingua* 74: 141–166.
Rizzi, Luigi (1990). *Relativized Minimality.* MIT Press, Cambridge, Ma.
Roberts, Ian (1991). Excorporation and Minimality, *Linguistic Inquiry* 22.1: 209–218.
Ross, John Robert (1967). *Constraints on Variables in Syntax.* PhD dissertation, MIT.
Steiner, Richard (1997). Ancient Hebrew, in *The Semitic Languages*, ed. by Robert Hetzron, 145–173. Routledge, London.
Svenonius, Peter (1996). The optionality of particle shift, *Working Papers in Scandinavian Syntax* 57: 47–75.
Taraldsen, Tarald (1998). V-movement and VP-movement in derivations leading to VO order, paper presented at VO and OV workshop in Tromsø.
Zwart, Jan Wouter (1993). *Dutch Syntax: A Minimalist Approach.* PhD dissertation, University of Groningen. Grodil, Groningen.

THE COPY THEORY OF MOVEMENT AND LINEARIZATION OF CHAINS IN THE MINIMALIST PROGRAM

by *Jairo Nunes*
reviewed by *Hans-Martin Gärtner*

Summary by *the author*

A considerable amount of research within the Principles and Parameters Theory has been devoted to properly characterize the properties of movement, traces, and chains. In the recent developments of the Principles and Parameters Theory which have culminated with the proposal of a Minimalist Program for linguistic theory (see Chomsky 1995), these issues arise anew in face of the elimination of much of the rich theoretical apparatus previously available.

Chomsky (1993) incorporates into the Minimalist Program the "copy theory of movement", according to which a trace is a copy of the moved element which is deleted in the phonological component, but remains available for interpretation at LF. Under this view, the operation Move is a complex operation comprised of (at least) three suboperations (see Chomsky 1993: 22; 1994: fn. 13; 1995: 250): (i) a suboperation of copying; (ii) a suboperation of merger; and (iii) a suboperation of trace deletion. In addition, Move should be followed by an operation of chain formation relating the relevant copies.

There are several conceptual inadequacies in this picture. First, if no explanation for why "lower" copies must be deleted in the phonological component is provided, the notion of a trace as a primitive is reintroduced. To put it more generally, the simplest and, therefore, most desirable version of the copy theory of movement should take traces and heads of chains to be subject to the same principles and be accessible to the same operations. Any difference between heads of chains and traces, such as phonetic realization, for instance, should follow from independently motivated properties of the computational system, rather than being idiosyncratic properties of the chain links themselves.

Deletion of traces (lower chain links) becomes even more enigmatic, if we adopt the core Minimalist assumption that economy considerations play a role in determining the set of admissible derivations in a given language or universally. Consider for instance the structure in (1) below, where *John* has moved to the subject position and left a copy behind. The derivation of (2a) from (1) requires one application of deletion targeting the lower copy of *John*, apparently being less economical than the derivation of (2b), which involves no application of deletion. Thus, if the derivations of (2a) and (2b) were to be compared in terms of economy, we would wrongly predict that the derivation of (2b) should rule out the derivation of (2a).

(1)
[John [was [arrested John]]]

(2)
a. John was arrested.
b. *John was arrested John.

Another conceptual problem with the computational system as proposed in Chomsky (1994, 1995: chap. 4) is that Merge is taken to be an operation in its own right in certain cases, and a suboperation (of Move) in other cases. In an optimal system, we should in principle expect Merge to have the same theoretical status in every computation. Finally, as is emphasized by Brody (1995), if chain formation and Move express the same type of relation, a theory which contains both notions is redundant.

This dissertation develops a strictly Minimalist version of the copy theory of movement which overcomes the conceptual problems raised above and has a broader empirical coverage than the versions developed in Chomsky (1993, 1994, 1995: chap. 4). It proposes that the fact that a chain cannot have more than only link overtly realized (see (2b)) follows from Kayne's (1994) Linear Correspondence Axiom (LCA), according to which the linear order of a PF sequence is determined by asymmetric c-command. Under the assumption that the two copies of John in (1) are "nondistinct" (they relate to the same element in the initial numeration), no linear order can be established in accordance with the LCA. Given that the verb *was* in (1), for instance, asymmetrically c-commands and is asymmetrically c-commanded by the "same" element, namely *John*, the LCA should require that *was* precede and be preceded by *John*, violating the asymmetry condition on linear order. Put simply, deletion of all but one link is forced upon a given chain CH in order for the structure containing CH to be linearized in accordance with the LCA. The derivations of (2a) and (2b) therefore cannot be compared for economy purposes, because only the former yields a PF object.

The next question to be addressed then is why it is the case that only traces are deleted for purposes of linearization, but not heads of chains.

In other words, why does the structure in (1) surface as (3a) and not as (3b), given that both structures can be linearized in compliance with the LCA?

(3)
a. John was arrested.
b. * Was arrested John.

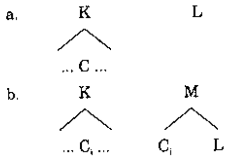
The basic idea is that heads of chains become "different" in the course of the derivation due to their participation in checking relations. For the sake of illustration, consider the Case-feature of *John* in the course of the derivation of (1), as shown in (4) below.

(4)
a. [was [arrested John-CASE]]
b. [John-CASE [was [arrested John-CASE]]]
c. [John [was [arrested John-CASE]]]

Assume that the Case-feature of the upper instance of *John* in (4b) is eliminated after being checked against the finite T head, as represented in (4c). Since Case-features are not PF objects, deletion of the lower copy of *John* in (4c) yields a legitimate PF object, whereas deletion of the upper copy does not; hence the contrast between (3a) and (3b). This rough idea is technically implemented in terms of economy considerations regarding the elimination of formal features in the phonological component.

As for the other conceptual problems mentioned above (the dual status of Merge and the redundancy between Move and chain formation), it is argued that Move is not a primitive operation of the computational system. It is rather a mere description of the interaction of the independent operations Copy, Merge, Chain Reduction (deletion of chain links for linearization purposes), and Form Chain. Thus, Merge is always an operation, and applications of Copy and Merge are not redundant with Form Chain. Empirical evidence for this approach is provided by instances of "sideward movement", where the computational system copies a given constituent C of a syntactic object K and merges C with a syntactic object L, which has been independently assembled and is unconnected to K, as illustrated in (5).

(5)



(7)–(11) below illustrate the relevant steps of a sideward movement analysis of the across-the-board extraction construction in (6) (irrelevant traces were omitted). The phrase *which book* is copied from K in (7a) and merges with L in (7b), yielding M in (8a). The auxiliary *did* is then copied from K and merges with M, yielding N in (8b). Further computations involve: (i) the merger of *Mary* with N, forming O in (9a); (ii) the merger of O and K, forming the syntactic object P in (9b); and (iii) the merger of the interrogative complementizer Q with P, yielding the syntactic object R in (9c). In R, the copies of *which book*, for example, cannot form a chain because one does not c-command the other. Thus, at this point of the derivation, "movement" and chain formation are not redundant. Since the complementizer Q in (9c) has strong features (see Chomsky 1995), additional copies of *did* and *which book* are made and merge with R, yielding S in (10). In (10), the highest copy of did forms a different chain with each of the lower copies; likewise, the highest copy of which book forms a different chain with each of the lower copies. Formation of each of these chains complies with the Minimal Link Condition (see Chomsky 1995); crucially, since neither of the relevant lower copies c-commands the other, they are equally close to the highest copy. Chain Reduction then deletes the traces (lower copies) of each of the four chains in the phonological component, yielding T in (11), which is finally realized as (6).

(6)
Which book did Mary buy and John read?

(7)
a. K = [and [John did_j read [which book $]_i$]
b. L = buy

(8)
a. M = [buy [which book $]_i$]
b. N = [did_j [buy [which book $]_i$]]

(9)
a. O = [Mary [did_j buy [which book $]_i$]]]
b. P = [Mary did_j buy [which book $]_i$] [and [John did_j read [which book $]_j$]]
c. R = [Q [Mary did_j buy [which book $]_i$] [and [John did_j read [which book $]_i$]]]

(10)
S = [[which book $]_i$ did_j +Q [Mary did_j buy [which book $]_i$] [and [John did_j read [which book $]_i$]

(11)
T = [[which book $]_i$ did_j +Q [Mary buy] [and [John read]]]

Parasitic gap constructions also receive a straightforward analysis in terms of sideward movement. In particular, it is shown that Chain Reduction constrains sideward movement in such a way that it makes possible to subsume some of the core properties of parasitic gap licensing such as the S-Structure condition and the anti-c-command requirement to properties of standard movement.

This attempt to eliminate traces as primitives and to break down Move into more basic independent operations has large implications for syntactic theory in general and for the Minimalist Program in particular, as discussed throughout the dissertation.

References

Brody, Michael (1995). *Lexico-logical form.* MIT Press, Cambridge, Mass.
Chomsky, Noam (1993). A minimalist program for linguistic theory. In *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 1–52. MIT Press, Cambridge, Mass.
Chomsky, Noam (1994). Bare phrase structure, *MIT Occasional Papers in Linguistics* 5, Cambridge, Mass.
Chomsky, Noam (1995). *The Minimalist Program*, MIT Press, Cambridge, Mass.
Kayne, Richard (1994). *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.

Review
by *Hans-Martin Gärtner*

Introduction

This dissertation can doubtlessly be called one of the most exciting and thought-provoking pieces of work written within the minimalist framework. First, it provides a lucid and comprehensive introduction to minimalist syntax. Secondly, movement theory is reanalyzed in a way that raises — and shows an interesting way to answer — the following puzzling questions:

Why is it that only one member of a syntactic chain gets phonetically realized? Why is that member invariably the hierarchically most prominent one reaching PF?

Finally, the technically intricate revision of movement theory creates some rather surprising spin-off insofar as ATB and parasitic gap constructions receive a natural syntactic analysis, the core constraints on the latter falling out directly.

These three achievements roughly correspond to chapters II, III, and IV of the dissertation respectively. (Chapter I gives a brief introduction and chapter V concludes the thesis.) Following a brief note on chapter II, this review will concentrate on chapters III and IV.

Chapter II ("Theoretical Framework") devotes 176 pages to laying out and critically analyzing the essentials of minimalist syntax, adjusting whatever the later constructive chapters demand. For instance, one of the technically more sophisticated maneuvers is the elimination of Procrastinate. Likewise, in this chapter, Nunes shows how we can dispense with the Uniformity Condition on chains, which Chomsky (1995) employed in capturing the Structure Preservation Hypothesis within his Bare Phrase Structure model. What looks somewhat awkward is Nunes's proposal to put strong features, which are, after all, the driving force of overt movement operations, into the set of **phonological** features of the lexical items bearing them. Given that strong features have to be checked in syntax, this conflicts with the otherwise fairly plausible (and restrictive) assumption (1).

(1)
[O]vert operations cannot detect phonological features at all.
(Chomsky 1995a, p.230)

Copy + Merge Theory of Movement

Let's now turn to the core issue of Nunes's dissertation, the "Copy + Merge Theory of Movement". As is well known, minimalist syntax relies on two structure-building operations, namely, *Merge* and *Move*. Merge transforms two (unordered) trees T1 and T2 into a single (unordered) tree T3 composed of T1 and T2, as exemplified in (2). (Take all trees/subtrees to be unordered, unless stated otherwise. I use traditional notation in order to shortcut issues tangential to this review.)

(2)
Merge(T1 = [arrest], T2 = [John]) ⇒ T3 = [arrest John]

Several conditions can be — and usually are — imposed on Merge. Thus, T1 and T2 can be required to become (proper) subtrees of T3 in the technical sense. This is usually formulated as a cyclicity condition or ban on 'internal' Merge. Secondly, the operation Move, later called "Attract" (Chomsky 1995a, chapter 4), transforms a single tree T1 into a tree T2 in the way specified by the following suboperations:

(i) locate two subtrees T3 and T4 of T1, T3, the attractor, c-commanding T4, the attracted element;
(ii) verify that T3 and T4 contain features able to establish a checking relation; (This is the Last Resort requirement. It could also be part of step (i), assuming that T3 only 'attracts' the right kind of object.)
(iii) verify that there is no subtree T5 of T1 closer to T3 than T4, such that T3 and T5 could establish a checking relation equivalent to the one specified in (ii);
(iv) copy T4;
(v) merge T4', the copy of T4, into the checking domain of T3;
(vi) form a chain between T4' and T4;
(vii) check the features in an established checking relation between T3 and T4' (and T4?); (The ordering of (vi) and (vii) as well as the uncertain status of T4 with respect to (vii) raise non-trivial questions.)
(viii) mark T4 as phonetically unrealizable. (This step is trivial in the covert component).

A much simplified example of Move is given in (3). (//... // encodes the PF condition on T4.)

(3)
Move (T1 = [was arrested John]) ⇒ T2 = [$John_i$ [was arrested //$John_i$//]]

One of Nunes's major claims is that (i)–(viii) should not be seen as a genuine compound (cf. Collins 1997). Rather, Copy, (in (iv)) should be allowed to operate freely and 'blindly'. Merge (in (v)) is an independent operation anyway, as characterized above. (ii), (iii), and (part of) (i) can be integrated into an independent Form Chain operation, which means that (vi) will be a separate and expanded operation. (There are various options as to where exactly (vii) should go, which I sidestep.) Crucially, (viii) is nothing to worry about at all outside the PF-branch of the computational system. Copies spelled-out into the PF-branch will all be phonetically realizable.

The question that comes up is how such a system gets rid of the superfluous copy/copies at PF? The answer, which makes up the bulk of chapter III ("Linearization of Nontrivial Chains at PF (or Why Traces are not Phonetically Realized)"), has to do with Kayne's (1994) Linear Correspondence Axiom (LCA). The LCA roughly states that asymmetric c-command between non-terminal nodes maps into linear precedence among the terminals dominated by the respective non-terminals. This principle is adapted by Nunes in the following way.

(4) Linear Correspondence Axiom (LCA)
A category α precedes a category β iff:
(i) α asymmetrically c-commands β; or
(ii) γ precedes β, and γ dominates α. (p.177)

For the sake of simplicity, we take categories to correspond to subtrees. At PF, the complex operation **Linearize** converts unordered trees into ordered ones on the basis of (4). Precedence among subtrees as it emerges from (4) is isomorphic to the familiar precedence relation among nodes of constituent structure trees (cf. Partee et al. 1993, p. 441), although exceptions arise where the correspondence between categories and subtrees breaks down and where c-command is undefined. Like precedence among nodes, precedence among subtrees is required to be a strict partial order (transitive, irreflexive, and asymmetric). Crucially, this property is taken to be violated by structures that contain multiple copies of the same subtree, such as (5).

(5)
[$_{(IP)}$ John$_i$ [was [$_{(VP)}$ arrested John$_i$]]]

Precedence among subtrees will contain both <John$_i$, was> and <was, John$_i$>, because the occurrence of *John$_i$* in SpecIP asymmetrically c-commands *was* and *was* asymmetrically c-commands the occurrence of *John$_i$* in VP (p. 273). Thus, precedence among subtrees, not being asymmetric, is not well-formed and Linearize cannot proceed, that is, the derivation does not converge. Note that, in order to derive such a result, asymmetric c-command must be able to distinguish occurrences of *John$_i$*, while precedence among subtrees isn't. Slightly paradoxically <John$_i$, John$_i$> must be a member of asymmetric c-command in (5). It follows that precedence among subtrees will be neither asymmetric nor irreflexive (p. 273). These subtle distinctions, of course, have to be based on clear criteria of identity together with an exact specification of the structures with respect to which ordering properties are verified. The former requirement is meant to be met by the following assumption.

"[T]ake a term of a given structure to be inherently distinct from all the other terms of that structure unless it is specified as being a copy. In other words, the output of a copying operation targeting a term T is a term T', which is interpreted at the C-I interface as nondistinct from T or other copies of T and distinct from every other term by default. The copying operation can be informally described as follows: (i) if a term T has no index, Copy targets T, assigns an unused index *i* to it and creates a copy of the indexed term; or (ii) if a term T was already indexed by a previous copying operation, Copy simply creates a copy of the indexed term." (p. 86)

As with categories before, we can take terms to correspond to subtrees unless stated otherwise. Crucially, coindexation makes subtrees *nondistinct*, a property geared to the purpose of treating two objects as one and the same when it comes to interpretation. Reference to the C-I interface is at least suggestive of that. Analogously, Linearize is close enough to the A-P interface to call it an 'interpretive' operation, so construction of precedence among subtrees as part of Linearize can be considered sensitive to nondistinctness, while verification of asymmetric c-command isn't.

Thus far it looks as if copying before Spell-Out inevitably leads to non-wellformedness at PF, an absurd consequence if movement is to be captured by Copy + Merge. But before we see how Nunes turns such an apparent vice into a virtue, let's have a closer look at Linearize.

Clearly, precedence among subtrees is not a linear order in the technical sense, that is, it is a *partial* not a *total* order. Precedence according to (4) will not be defined of subtrees/categories ordered vertically, i.e., by dominance. Only precedence among terminals can be required to be a linear order, derivable from the *locally* linear order induced on non-terminals by asymmetric c-command (Kayne 1994, p. 4). Nunes is not always terminologically careful enough to mark these differences. Yet, apart from open issues related to the absence of a traditional distinction between terminals and non-terminals and the ban on non-branching projection in minimalist bare phrase structure (for discussion see section II.17 and Chomsky 1995b), there seems to be good reason for Nunes to have Linearize construct precedence among subtrees and not just precedence among terminals. Consider example (6).

(6)
[$_{(IP)}$ [$_{(DP)}$ the boy]$_i$ [was [$_{(VP)}$ arrested [$_{(DP)}$ the boy]$_i$]]]

Copying has applied to the complex DP, so that's where the index went. If we take X°-heads as terminals, then precedence among terminals of (6) induced by asymmetric c-command via definition (4) would contain pairs such as <boy, was> and <was, boy>. Without indices it is unclear how to treat the two occurrences of *boy*. They are definitely not identical in the sense of Leibniz's principle, according to which "if x and y are identical, then y has every property of x" (Smullyan 1971, p. 35). One occurrence is part of the specifier of IP while the other is part of the complement of V. Does this mean that default distinctness applies, as invoked earlier in the case of occurrences of subtrees with identical featural content? That would depend on the exact implementation of such a default. (It is doubtful, by the way, whether anything but another layer of indices can do that job in the set-based representations of bare phrase structure.) Thus, unless indices percolate to each subtree of a subtree copied, precedence among subtrees is the proper place to rule out (6) as it stands. Obviously, precedence among subtrees will contain <[(DP) the boy]$_i$, was> and <was, [(DP) the boy]$_i$>. Coindexation means nondistinctness, and the latter leads to a violation of the asymmetry and irreflexivity requirement put on precedence among subtrees. Linearize cannot proceed and (6) is ruled out on a par with (5). (Just after finishing this review, Nunes informed me (p.c.) that it is index percolation he would favor in dealing with cases like (6). So, things attributed here to precedence among subtrees will have to be worked out for precedence among terminals, *mutatis mutandis*.)

Let's postpone inspecting the solution to the already mentioned apparent impasse created by (5) and (6) for yet another round and consider a conceptually interesting issue first.

LCA and the Minimalist program

Kayne's LCA is a powerful condition that almost completely derives the form phrase-structure representations of natural language sentences should have. It is doubtful, whether the highly derivational Chomskyan system, by and large adopted by Nunes, really has to resort to the full power of the LCA, if the sole aim is to achieve PF-linearization. Note first that Kayne does not seem to take copies of movement to lead to LCA-violations (Kayne 1994, p. 133, fn. 3). But, how else could reflexive or symmetric pairs enter precedence among subtrees or precedence among terminals? The former case clearly requires that some subtree asymmetrically c-command itself. Such a thing could result from relaxing the "Single Mother Condition" (7) on dominance (N = node, ID = immediately dominates; cf. Blevins 1990, p. 48).

(7) Single Mother Condition
$(\forall x,y,z \in N) [(xIDz \wedge yIDz) \rightarrow x = y]$

A subtree T1 allowed to be immediately dominated by two distinct nodes x and y, such that either x dominates y or y dominates x, should yield <T1, T1> as a member of asymmetric c-command. However, since the same subtree cannot undergo Merge twice in the Chomskyan system, that is, it cannot acquire two mother nodes, such a violation simply cannot arise. The irreflexivity condition, thus, holds trivially. Next, violations of asymmetry would either follow from the same kind of relaxation of (7), namely, if there is a distinct node z dominance-wise between x and y, such that z immediately dominates an additional subtree T2. T2 could then both asymmetrically c-command and be asymmetrically c-commanded by T1. But, this is a non-issue as we've just seen. Alternatively, two subtrees could symmetrically c-command each other. Crucially, the latter case has been obviated in the minimalist system by stipulation. Nunes's version of this hinges on the *inaccessibility* of intermediate projections and lower segments of adjunction structures.

(8) C-Command
Where α and β are accessible to the computational system, α c-commands β iff:
(i) α does not dominate β;
(ii) α ≠ β;
(iii) every category dominating α dominates β. (p.177)

Sidestepping further complications, we can say that whenever subtree T1 adjoins to subtree T2, the resulting sister of T1 will not be *accessible* any longer. T1 can asymmetrically c-command into its sister and therefore precede every subtree dominated by its sister. Likewise, whenever two complex subtrees undergo 'non-adjunctive' Merge, the one that projects won't be accessible any longer, so the non-projecting one can asymmetrically c-command into its projecting sister and precedence among the respective sets of subtrees is fixed. Thus, except for the case of two terminal subtrees in a sisterhood relation — for which an extra assumption is required because, as already mentioned, nonbranching projection isn't allowed — no violation of asymmetry can arise. Again, compliance with that constraint has been built into the system and need not be invoked as an extra condition on representations.

We could, for the sake of argument, allow the LCA to do some genuine work in filtering out specifiers in the standard sense (i.e. sisters of intermediate projections) and allowing only adjoined 'specifiers' along the lines proposed by Kayne (1994). Intermediate projections would then have to be accessible with respect to c-command as defined in (8). Structures like (9) would then give rise to an asymmetry violation.

(9)
[$_{(XP)}$ [$_{(YP)}$ the cat] [$_{(X')}$ is [$_{(ZP)}$ on the mat]]]

Given that among other things X' asymmetrically c-commands *cat* and that YP asymmetrically c-commands *is*, both <cat, is> and <is, cat> should hold of precedence among subtrees or precedence among terminals. This violates the LCA. It is unclear, however, how to represent and exclude this in Nunes's system. Once again, indices, being reserved for copies, wouldn't apply here. Thus, the case in which strictly identical subtrees (i.e. subtrees to which Leibniz's principle does apply) get into conflict with the asymmetry condition on linearization doesn't receive as straightforward a treatment as one would expect. (An attempt to derive the ban on n-ary branching structures, for n ≥ 3, would be defective in the same way.) But if the canonical cases the representational LCA condition was meant to capture are not an issue in a theory invoking the LCA, this is likely to cast some doubt on how 'natural' the LCA-based approach to the problem caused by stipulated quasi-identical ("nondistinct") subtrees, i.e. copies/traces, really is. There surely is a very creative idea behind all of that, but its implementation doesn't come for free.

In fact, minimalist systems could employ some very simple top-down algorithm for converting unordered trees into ordered trees. Hitting on a pair of sister-subtrees T1 and T2, plus and minus accessible respectively, let T1 precede T2. Hitting on a pair of sister-subtrees, both accessible, let the one that is an X°-head precede the other subtree. This is all there has to be done (modulo the already mentioned extra assumptions concerning X°-sisters at the bottom of the tree). Most importantly, never does the system have to keep track of the content of any subtrees throughout linearization. This would be necessary if one were expecting one and the same subtree to (illicitly) appear twice in the resulting ordered tree, an impossibility as far as the ones that obey Leibniz's principle are concerned. For Nunes's solution to work, however, there has to be an additional check of the resulting ordered tree in order to detect nondistinct subtrees. This would seem to amount to pair-wise inspecting a full chart of precedence among subtrees (or precedence among terminals), on the size of which, as is well-known, there is no upper bound.

I have been pushing this rather abstract point at the risk of using formulations the author doesn't fully subscribe to. It is my strong impression, though, that the delicate issue of syntactic individuals and how they are structurally individuated affects Nunes's approach in a potentially adverse fashion, even if different linearization algorithms make this look more or less transparent.

Let's now return to the question why (5) ultimately gets realized as (10) at PF.

(10)
John was arrested

This is due to the possibility of **Chain Reduction** operating in the PF-component before Linearize applies.

(11) Chain Reduction
Delete the minimal number of terms of a nontrivial chain CH which suffices for CH to be mapped into a linear order in accordance with the LCA (p. 315)

Suppose the two occurrences of $John_i$ have formed a chain in (5). Since each copy consists of one subtree, deleting any one of the two will allow compliance with the LCA. Asymmetric c-command as computed over the thus reduced tree structures will not produce any offensive pairs involving nondistinct copies. Therefore, precedence among subtrees will be a strict partial order as required. Alongside of (10), (12) would be another well-formed output of Linearize.

(12)
was arrested John

Even the option of deleting both occurrences of $John_i$, leads to an LCA compatible structure. However, Nunes assumes that transderivational economy blocks the application of two deletions where one such operation would likewise have produced a convergent output. Thus, (13) is an impossible representation of (5).

(13)
was arrested

In sum, Nunes gives the following answer to the first of the fundamental questions he raised. Phonetically realizing more than one member of a nontrivial chain leads into conflict with the LCA. Phonetically realizing no member of a nontrivial chain is blocked by transderivational economy, which prevents Chain Reduction from applying more often than necessary. The survival of one member is compatible with linearization to proceed, so deleting that member is a superfluous step.

On a more conceptual note, it must be said that invoking Chain Reduction instead of just 'Reduction' pays tribute to the fact that, even if Move is dismantled as a unified operation, the effects standardly attributed to it have to be dealt with in a special way, at least to some extent.

FF-elimination

Nunes (section III.7.1) addresses "*wh*-Copying", i.e. cases that appear to spell-out more than one copy of a *wh*-chain, such as the German example given in (14).

(14)
Wo glaubst Du wo Maria wohnt
Where think you where Mary lives
'Where do you think Mary lives?'

The answer to the second fundamental question, namely, the question as to which member of a chain gets realized can be understood as soon as we've found out why the system prefers (10) to (12), repeated as (15a)/(15b) respectively.

(15)
a. John was arrested
b. was arrested John

Assuming that Linearize ultimately produces a string of X°-heads (p. 274), there is still another step to take before a phonetic representation π can be constructed. (From the perspective of prosodic phonology it is far from clear that a string of terminals is sufficient. Rather, it would seem, that mapping into prosodic constituents has to start from ordered syntactic constituents, so the strategy of constructing and evaluating precedence among subtrees rather than precedence among terminals as discussed above would be vindicated on these grounds.) The X°-elements still contain formal features. These will be eliminated on an item-by-item basis in accordance with (16).

(16) Formal Feature Elimination (FF-Elimination)
Given the sequence of pairs σ = <(F,P)1, (F,P)2, ..., (F,P)n> such that σ 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 σ to satisfy Full Interpretation at PF. (p. 315)

Crucially, (15a) and (15b) differ in that the subtree *John* in the former already lacks its Case-feature, due to having checked it off in the checking domain of I°. This saves one application of FF-Elimination. The subtree *John* in VP of (15b), however, has its Case-feature still unchecked. Thus, FF-Elimination will have to apply at least one time more often. Other things being equal, transderivational economy therefore chooses (15a) and blocks (15b). It is quite easy to see now why the hierarchically most prominent member of a chain reaching the PF-branch gets realized. Movement is driven by feature checking. The functional projections dominate the lexical ones. And, there is no lowering. The three instances of deletion, Checking, Chain Reduction, and FF-Elimination will interact in the desired fashion if it is exactly the hierarchically most prominent copy reaching PF that has fewer features due to checking. Chain Reduction targeting entire subtrees doesn't care about the number of features present in the chain members deleted. But FF-Elimination makes a difference. Thus, the more features can be gotten rid of via Chain Reduction the fewer times FF-Elimination has to apply. This logic favors the survival of the copy that has undergone a checking operation earlier, i.e., the hierarchically most prominent one.

Further adjustments are nevertheless required to deal with ECM constructions (17a) and *wh*-movement (17b), as well as the general question of successive-cyclic movement (17c).

(17)
a. We expected $[_{IP}$ $John_i$ to be $[_{VP}$ arrested $John_i$]]
b. $[_{CP}$ Who_i did $[_{IP}$ they $[_{VP}$ arrest who_i]]]
c. $[_{CP}$ Who_i do you think $[_{CP}$ who_i $[_{IP}$ they $[_{VP}$ arrested who_i]]]]

If movement of $John_i$ in (17a) and who_i in (17b) only checks interpretable features, the two nondistinct copies still carry the same number of formal features at PF. FF-Elimination should optionally allow (18a) and (18b) to be well-formed.

(18)
a. We expected to be arrested John
b. Did they arrest who

This problem is dealt with by the assumption that

"when participating in an overt checking relation, a [+interpretable] feature can optionally be checked with respect to PF, becoming invisible at this level." (p. 307)

Invisibility means that FF-Elimination doesn't have to deal with such a feature, that is, the D-feature of $John_i$ in SpecIP of (17a) and the D/*wh*-feature of who_i in SpecCP of (17b). Optionality of checking [+ interpretable] features with respect to PF is required to properly handle successive cyclic movement (17c). If who_i in the intermediate SpecCP were to check its D/*wh*-feature, the final step into the matrix SpecCP would, in the best case, produce a copy that contains equally few PF-relevant features as the one in the intermediate position. Thus, (19) should become an optional realization of (17c).

(19)
Do you think who they arrested

To prevent this, it is assumed that

"a feature which has been made invisible at either interface level by a checking operation cannot participate in any further checking relation." (p. 302)

Thus, the Q-feature of the matrix C° would have to remain unchecked if the option of checking the D/*wh*-feature of who_i in the intermediate SpecCP had been chosen. Of course, the theory has to be further strengthened in order for (18) not to remain an option. This is done in the following way.

"[E]conomy considerations dictate that two elements in an overt checking relation should have the greatest number of features checked with respect to PF (up to convergence)." (p. 307)

The D and D/*wh*-features of $John_i$ and who_i in (17a)/(17b) *can* be checked in SpecIP and SpecCP respectively without causing non-convergence, so they *must* be checked. (18) is therefore ruled out. I note in passing but refrain from discussing the heavy reliance of this system on transderivational economy. (On this issue see Chomsky 1996, Johnson & Lappin 1997, and Wilder et al. 1997).

Interestingly, Nunes's approach shows a straightforward way out of the puzzle created by local feature checking and copying. Standard checking after movement will involve configurations like (20).

(20)
$[\,[\ldots F^{*}\ldots]_i \;\; {}^{*}F\;[\ldots\ldots(copy[\ldots F^{*}\ldots]_i)\ldots]\,]$

A checkable feature F* reaches the checking domain of a functor *F, able to check against F*. At the same time, a copy of F* is left in situ. If the locality condition on checking is taken in its strictest sense, the lower copy of F* must not be affected when the higher one checks against *F. In such a case, however, the computation will never exhaust the resources driving movement operations (cf. Stabler 1996). Chomsky proposed that

"the simplest assumption is that the features of a chain are considered a unit: if one is affected by an operation, all are." (1995a, p. 381, fn. 12)

Nunes, on the other hand, capitalizes on leaving the lower copy of F* untouched. The imbalance checking creates among copies of a chain is then put to use at PF to determine the member of a chain to be phonetically realized (as we have seen). Getting rid of the unchecked feature at LF, however, requires *Chain Uniformization*, another deletion operation.

ATB and parasitic gaps

Let us now turn to the surprising spin-off yielded by Nunes's construal of movement as Copy and Merge (Chapter IV: "Sideward Movement and Linearization of Chains at PF"). If Copy is allowed to apply independently of movement, one would expect coindexed ('nondistinct') copies of one and the same subtree to occur in positions not related by the defining properties of Move, that is, c-command, Last Resort, and the Minimal Link Condition (MLC). Nunes dissociates these properties from Copy and Merge. They constitute the independent operation **Form Chain**.

(21) Form Chain
The syntactic objects α and β can form the nontrivial chain CH = (α, β) only if:
(i) α c-commands β;
(ii) α is nondistinct from β;
(iii) (at least) one given feature F of α enters into a checking relation with a sublabel of K, where K is the head (of the projection) with which α merges, and the corresponding feature F of β could enter into a checking relation with a sublabel of K; and
(iv) there is no syntactic object γ such that γ has a feature F' which is of the same type as the feature F of α, and γ is closer to α than β is. (p. 424)

According to Nunes, such a dissociation is exactly what we find in ATB and parasitic gap constructions. Let's just consider the latter.

(22)
a. Which article did you file without reading
b. [[which article]$_i$ did you [$_{(VP1)}$ file [which article]$_i$] [without [$_{(VP2)}$ reading [which article]$_i$]]]

The copies of *which article* in VP1 and VP2 (22b) cannot form a chain since they cannot satisfy the c-command and Last Resort requirement. The copy of *which article* in SpecCP, however, can form one chain with each of the in situ copies. Two applications of Chain Reduction will then yield (22a). This already highlights the major constraint on "sideward movement", i.e. the licensing of nondistinct copies that cannot form a chain. Unless they acquire some independent antecedent before Spell-Out, the LCA will be violated. Thus, there is no LF-licensing of parasitic gaps, as (23) illustrates.

(23)
a. * Who praised which article without reading
b. Who praised [which article]$_i$ [without reading [which article]$_i$]

The pairs <[which article]$_i$, without> and <without, [which article]$_i$> (among many others) will enter precedence among subtrees, so that Linearize is blocked. Obviously, LF-raising cannot affect such a PF-property. S-structure licensing of parasitic gaps is thus predicted by the theory without extra assumptions. Likewise, another major constraint, namely, the ban on c-command between the primary and the parasitic gap, falls out directly. Consider (24):

(24)
a. * Who called you before you met
b. Who$_i$ [who$_i$ called you [before you met who$_i$]]

Assuming the 'primary' copy to c-command the 'parasitic' one, these two cannot form a chain, because that chain doesn't meet the Last Resort requirement. Neither of the two enters any checking relation. At the same time, Form Chain between the copy in operator position and the 'parasitic' one violates the MLC, given that the 'primary' copy intervenes. But, if the 'parasitic' copy cannot enter any chain, the familiar LCA violation results. Again no extra assumptions are necessary to derive this. Structures like (24) reveal another subtlety of Chain Reduction. (24b) could be saved by deleting both copies of the 'primary' chain, which would yield PF-output (25).

(25)
Did you file without reading which article

This is prevented by conditioning Chain Reduction (11) on linearizability of just the chain being reduced, irrespective of whether the entire structure will eventually satisfy the LCA. Thus, the linearization algorithm has to be even more fine-grained in being able to keep track of which chain subtrees belong to.

Two further issues come to mind immediately in connection with Nunes's treatment of parasitic gaps that need to be explored. A theory heavily relying on c-command as required by the LCA must also be very explicit about the exact structural positions of extraposed constituents, if for example the difference between (22) and (24) is to be captured. Rightward adjunction not being an option, the author proposes a coordination-like treatment of adjuncts, the full detail of which would have to be worked out. Secondly, one has to worry about cases of parasitic gaps for which the Anti-Command condition doesn't seem to hold. This issue has recently been brought up again by Brody (1995) in connection with Hungarian examples like (26).

(26)
Kiket$_i$ szeretnél t'$_i$ ha t$_i$ eljönnének anékül hogy meghivtál volna e$_i$
Who would you like if came without you having invited
(Brody 1995, p.84)

Remnant movement

Let me finally discuss the phenomenon of *Remnant Movement* which leads to "sideward movement" of a more complex kind, as (27) shows (cf. Müller 1997).

(27)
a. Elected by the committee John never was
b. [[elected [John]$_i$ by the committee]$_j$ [[John]$_i$ never was [elected [John]$_i$ by the committee]$_j$]]

Chains can be formed between *John$_i$* in SpecIP and *John$_i$* inside the VP in situ, as well as between the fronted VP and the VP in situ. Copying VP, however, has produced another copy of *John$_i$* inside the fronted VP. For lack of c-command, that occurrence of *John$_i$* cannot undergo Form Chain and thus won't be deletable under any kind of Chain Reduction. Consequently, (20a) cannot be derived in the system as it stands.

In sum

Summing up, it must be emphasized again that in spite of the various weaknesses pointed out, this dissertation covers an admirably broad range of syntactic issues, creating deductively nontrivial links between seemingly disparate phenomena.

References

Brody, Michael (1995). *Lexico-Logical Form*. Cambridge Mass.: MIT Press.
Blevins, James (1990). *Syntactic Complexity*. Doctoral Dissertation, University of Massachusetts at Amherst.
Chomsky, Noam (1995a). *The Minimalist Program*. Cambridge Mass.: MIT Press.
Chomsky, Noam (1995b). Bare Phrase Structure. In: Webelhuth, Gert (ed.), *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell, 383–439.
Chomsky, Noam (1996). Some Observations on Economy in Generative Grammar. Ms. MIT.
Collins, Chris (1997). *Local Economy*. Cambridge Mass.: MIT Press.
Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge Mass.: MIT Press.
Johnson, David & Shalom Lappin (1997). A Critique of the Minimalist Program. *Linguistics and Philosophy* 20.3: 273–333.
Müller, Gereon (1997). *Incomplete category fronting*. Dordrecht: Kluwer.
Partee, Barbara, Alice ter Meulen & Robert Wall (1993). *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.
Smullyan, Arthur (1971). Modality and Description. In: Linsky, Leonard (ed.), *Reference and Modality*. London: Oxford University Press, 35–43.
Stabler, Edward (1996) Acquiring and Parsing Languages with Movement. Ms. UCLA.
Wilder, Chris, Hans-Martin Gärtner & Manfred Bierwisch (eds.) (1997). *The Role of Economy Principles in Linguistics*. Berlin: Akademie-Verlag.

LEFT WITH THE SOUNDS

by *Bonny Sands* reviewing the CD Extinct South African Khoisan Languages

"Know me through my tongue", said Mukalap, one of the last fluent speakers of !Ora, in 1938, to an international meeting of phoneticians. The recording featuring his words, spoken in !Ora are especially poignant. If linguists forget about a language, who else will remember it?

Prof. Tony Traill of the University of the Witswatersrand had dug through the archives to find recorded examples of some of the "forgotten" languages of South Africa. As the leading expert on these languages, he translated and transcribed where it was possible, and resurrected voices from old, scratched wax cylinders.

A booklet accompanies the CD, and includes transcriptions, translations, pictures, and introductions to each of the 15 tracks. The pictures are as moving as the words. For instance, we see pictures of ‖Khaku and his family from 1936, and also pictures of two women who heard ǂKhomani as children, listening to ‖Khaku's words.

!Ora, ‖Xegwi, Ku |khaasi, ǂKhomani are now extinct, but thanks to this CD, we can hear them for the first time.

Track 2 is a good recording to use in introductory linguistics classes. With its IPA transcription, students can try to follow along with "The North Wind and the Sun" in ‖Xegwi. It is fun for students to hear a click language as it is spoken, with the clicks behaving 'just' like other consonants.

Most of the tracks do not lend themselves easily to analysis, but they all have something interesting to tell us. For instance, Track 3 has no accompanying transcription or even translation but its use in class may have a great impact on students. This recording, and an unpublished manuscript of notes by Robert Story, are all we know about the Ku |khaasi language. When students realize there are still many un- and under-described languages (including many Khoisan languages) in the world, it underscores part of the adventure of linguistics. The recording presents Ku |khaasi as an anomaly for classification, by having both bilabial clicks and uvular stops (characteristics of non-!Kwi, Southern Khoisan languages) and the presence of a word that looks cognate with 'child' in !Kwi languages.

All in all, the CD is a great deal of fun and should be a part of any Linguistics Department's collection. The only detraction that I can see is the high cost of the CD but one should keep in mind that the CD is being offered at cost.