BEYOND EXPLANATORY ADEQUACY

by

Noam Chomsky

&

ELIMINATING LABELS

by

Chris Collins

MIT Occasional Papers in Linguistics
Number 20

2001

Distributed by
MIT Working Papers in Linguistics
E39-245, MIT, Dept. of Linguistics
77 Massachusetts Avenue
Cambridge, MA 02139 USA
nutwpl@mit.edu // http://mit.edu/mitwpl/

Beyond Explanatory Adequacy*

Noam Chomsky

MIT, May 2001

I will assume here an approach to the study of language that takes the object of inquiry to be an internal property of persons, a subcomponent of (mostly) the brain that is dedicated specifically to language[1]: the human 'Taculty of Language" (FL), to adapt a traditional term to a new context. This "biolinguistic approach" was controversial when it took shape almost half a century ago, and remains so, but without warrant, in my opinion.[2] A stronger thesis is that the biolinguistic approach has a kind of privileged status, in that every constructive approach to human language and its use presupposes it, or something similar, at least tacitly. That too seems to me tenable, but I will not pursue the issue here.

FL appears to be a species property, close to uniform across a broad range. It has a genetically-determined initial state So, which determines the possible states it can assume. Suppose we also adopt — if it is too strong, only for convenience — a strong uniformity thesis for language acquisition that holds that each attainable state of FL is a further specification of So with parameters valued[3]; at So, all parameters are set with unmarked values. Then each attained state (including So) is a possible (I-) language L. We can then formulate without complication the familiar idealization: So (= LAD) maps'primary linguistic data (PLD) to L. We then seek to discover theories that meet the conditions of descriptive and explanatory adequacy — that are true, respectively, of L (particular grammars) and of So (universal grammar, UG).

It was clear from the outset that within the biolinguistic framework, this formulation is oversimplified in crucial respects. The initial conditions for (the abstract model of) language acquisition include more than So. The properties of the attained language L result from the interaction of three factors[4]:

(1) (1) individual experience (PLD), which selects among the options allowed by So
(2) So itself, a product of evolution
(3) general properties of organic systems

In this case the organic systems are computational systems that incorporate, it is reasonable to expect, principles of efficient computation. The picture is familiar in the general study of organic systems. Work of D'Arcy Thompson and Alan Turing on form and morphogenesis is a classic illustration of (3), but recognition of its importance goes back to the origins of evolutionary biology, and the basic point is a virtual truism: natural selection can only

---

* This paper will appear in the Proceedings of the Pontignano Conference, edited by L. Rizzi.
[1] As a system, that is; its elements might be recruited from, or used for, other functions.
[2] For some recent discussion, see Jenkins 2000, Chomsky 2000a, 2001.
[3] Or with the set of remaining choices narrowed in a "grammar competition model" of the kind investigated in Roeper 2000, Yang 2000; see the latter on antecedents and alternatives.
* See Chomsky 1965,59. Also Chomsky, Huybregts, and van Riemsdijk 1982,23.

function within a "channel" of options afforded by natural law, including properties of complex systems. One current example that may be suggestive in the present context is work by Christopher Chemiak, who has been exploring the idea that rninimization of "wire length," as in microchip design, should produce the "best of all possible brains," and has sought to explain in these terms the neuroanatomy of nematodes and some properties of nervous systems generally, such as the fact that the brain is as far forward as possible on the body axis.[5]

One can plausibly trace interest in factor (3) back to the Galilean intuition that "nature is perfect," from the tides to the flight of birds, and that it is the task of the scientist to discover in just what sense this is true. However obscure it may be, that intuition about what Haeckel called nature's "Sinn fur das Schone" has been a guiding theme of modern science ever since, perhaps its defining characteristic.[6]

In principle, then, we can seek a level of explanation deeper than explanatory adequacy, asking not only *what* the properties of language are, but *why* they are that way.

For familiar and substantial reasons, research concentrated on the problems of descriptive and explanatory adequacy, restricted to factors (1) and (2). The choice of terminology reflects the feeling that factors of category (3) were beyond the range of feasible inquiry, though they are always prominent even in pursuit of the narrower concerns, as part of the motivation for "best theory" considerations. That these limits might be transcended was suggested by the crystallization of the Principles & Parameters program in discussions at the Scuola Normale Superiore in Pisa in 1979 and subsequently. This approach offered a way to overcome the tension between descriptive and explanatory adequacy for the first time, and at once suggested that questions of category (3) might be directly addressed in a serious way. Various efforts were made through the 1980s, and a few years later attained some substantive results (e.g., irt recasting problems of "reconstruction" in terms of reduction of the theory of movement to its bare essentials) and seemed to show considerable promise, coming to be called "the minimalist program," by now with a rich and varied literature. Whether further optimism is warranted is hardly a topic for useful debate; as always in the case of research intuitions, time will tell.

Assuming that these questions can now be seriously placed on the research agenda, we can proceed further to disaggregate So into elements that have a principled explanation, and others that remain unexplained at this level of analysis, and must be attributed to something independent: perhaps path-dependent evolutionary processes, or properties of the brain that remain unknown. These would have to be studied along similar lines, sorting out the effect of general principles (physical, chemical, mathematical), interface conditions, and a residue to be accounted for in other terms. The principled elements of So are the conditions imposed on FL by the systems with which it interacts. If language is to be usable at all, its design must satisfy an "interface condition" IC: the information in the expressions generated by L must be accessible to other systems, including the sensorimotor (SM) and conceptual-intentional (C-I) systems that enter into thought and action. We can therefore restate the deeper "why-question." Insofar as properties of L can be accounted for in terms of IC and general properties of computational efficiency and the like, they have a principled explanation: we will have validated the Galilean intuition of perfection of nature in this domain.

The minimalist program is the attempt to explore these questions. Its task is to examine every device (principle, idea,...) that is employed in characterizing languages to determine to what extent it can be eliminated in favor of a principled account in terms of general conditions of computational efficiency and the interface condition that the organ must satisfy for it to function at all. Put differently, the goal is to determine just what aspects of the structure and use of language are specific to the language faculty, hence lacking principled explanation at this level.[7]

Summarizing, the initial conditions on language acquisition fall into the categories (i),

(2) (i) unexplained elements of So
(ii) IC (the principled part of So)
(iii) general properties

Principled explanation, going beyond explanatory adequacy, keeps to (ii) and (iii). An extremely strong minimalist thesis SMT - too much to expect — would be (3):

(3) (2i) is empty

Evidently, there are no a priori instructions about how to proceed on this path. The questions are empirical at every point, including the kinds of computational efficiency that FL selects; such elements of category (iii) are external to So -just as there is no gene that determines how particular proteins fold-, but unexplained (category (i)) elements of So may choose among these options. I will begin by adopting fairly standard assumptions (which are by no means innocent), proceeding along more controversial lines that seem to me reasonable and promising,[8] assuming the best case to hold unless the contrary is demonstrated, and putting to the side many attractive alternatives that are currently being pursued.

The language L generates a set of derivations. The last line of each derivation D is a pair <PHON, SEM>, where PHON is accessed by SM and SEM by C-I. Derivation D *converges* if PHON and SEM each satisfy IC; otherwise it *crashes* at one or the other interface. IC must be strong enough to allow sufficient diversity of 'legible" expressions at the SEM interface.

[5] Cherniak 1995. For an illuminating review of the state of the art, see Stewart 1998; briefer comments, Maynard Smith 1998. A more technical review appears in Maynard Smith et al. (1985). On Turing's contributions particularly, see Leiber 2001.

[6] For recent commentary relevant to linguistics, see Uriagereka 1998, Jenkins 2000, Martin and Uriagereka 2000, Freidin and Vergnaud 2001.

[7] The approach is complementary to others, for example, studies of language acquired under sensory deficit, or studies of other species that seek to identify properties that may be recruited for language but not specific to it. For suggestive recent results of these two categories, see Petitto et al. (2000), Ramus, et al. (2000).

[8] See Chomsky 2000b *(Minimalist Inquiries,* henceforth *MI),* 2001 *(Derivation by Phase,* henceforth *DbP),* presupposed in what follows, though with significant modification.

Exactly how this requirement should be formulated is not obvious. At least infinite legibility is presumably required, but that condition is not strong enough: it could be satisfied, for example, by indefinite reiteration of an element with no features that are uninterpretable at the interfaces, e.g., "No, No,...." A very strong condition would be that each derivation is "failure proof: there must be a way to extend it to a convergent derivation.' Though this condition may be too strong (see below), something like it has motivated much recent work, which has sought to eliminate comparison of derivations, backtracking and look-ahead, and "non-local" operations generally.[10]

So determines the set {F} of properties ("features") available for languages. Each L makes a one-time selection of a subset [F] of {F} and a one-time assembly of elements of [F] as its lexicon LEX, which we can take to be a classical "list of exceptions," putting aside further issues.[11] More controversially, for each derivation D, L makes a one-time selection of elements of LEX that will be accessed in D: a *lexical array* LA (a *numeration* if elements of LEX are accessed more than once). Each of these decisions involves a trade-off: memory requirements are restricted (massively, in the case of LA), but new concepts are introduced. Whether the decisions are correct is, as always, a (rather subtle) question of fact about language design. I will assume they are, but nothing here hinges directly on this.

Assume further that L has three components: *narrow syntax* (NS) maps LA to a derivation DNS; the *phonological component Q* maps DNS to PHON; the *semantic component* Z maps DNS to SEM. Z is assumed to be uniform for all L; NS is as well, if parameters can be restricted to LEX (as I will assume). <1> in contrast is highly variable among Ls. Optimally, mappings will satisfy the *inclusixxness condition,* introducing no new elements but only rearranging those of the domain. Assume this strong condition to be true of NS. It is surely not true of <I>, nor (on usual assumptions) of Z.[12]

Assume that all three components are cyclic, a very natural optimality requirement and fairly conventional. In the worst case, the three cycles are independent;[13] the best case is that there is a single cycle only. Assume that to be true. Then *9* and 2 apply to units constructed by NS, and the three components of the derivation of <PHON, SEM> proceed cyclically in parallel. L contains operations that transfer each unit to <t> and to Z. In the best case, these apply at the same stage of the cycle. Assume so. Then there is an operation TRANSFER, applying to the narrow-syjntactic derivation DNS:

(4) TRANSFER hands DNS over to * and to Z

We focus here primarily on the mapping to <1>, returning to their integration later: call it *Spell-Out,* S-O."

In this conception there is no LF: rather, the computation maps LA to <PHON, SEM> piece-by-piece cyclically. There are, therefore, no LF properties and no interpretation of LF, strictly speaking, though Z and $ interpret units that are part of something like LF in a non-cyclic conception.

Call the relevant units "phases".[15] It remains to determine what the phases are, and exactly how the operations work. I will assume, following *DbP,* that the phases are CP and *v?,* but crucially not TP, returning to some reasons.[16] When a phase is transferred to G>, it is converted to PHON. <I> proceeds in parallel with the NS derivation. O is greatly simplified if it can "forget about" what has been transferred to it at earlier phases; otherwise, the advantages of cyclic computation are lost. While the assumption may be somewhat too strong, let us assume it to be basically true, so that global properties of phonology (e.g., intonation contour) are superimposed on the outcome of cyclic operation of <C>.

Applied to a phase PH, S-O must be able to spell-out PH in full, or root clauses would never be spelled out.[17] But we know that S-0 cannot be *required* to spell-out PH in full, or displacement would never be possible. Consider a typical phase (5), with H its head:

(5)PH = [a [H p]]

Call a-H the *edge* of PH. It is a fact that elements of the edge may (or sometimes must) raise. A natural condition, which permits spell-out of root phrases and allows for meaningful cyclic computation, is that P must be spelled out at PH, but not the edge: that allows for head-raising, raising of predicate-internal subject to SPEC-T, and an "escape hatch" for successive-cyclic movement through the edge. Call this condition the "Phase Impenetrability Condition" PIC. However PIC is formulated exactly, it should have as a consequence that at the phase ZP containing phase HP:

(6) The domain of H is not accessible to operations, but only the edge of HP[18]

' See Frampton and Gutmann (2000).

[10] For discussion and references to other work on these particular topics (by Chris Collins, Sam Epstein, John Frampton and Sam Gutmann, Howard Lasnik, and others), see *MI, DbP.*

[11] Note that this convention entails that L changes when a new lexical item is introduced. Alternatively, LEX could be replaced by a generative system for constructing the possible lexical elements of L.

[12] These asymmetries allow for investigation of the internal structure of NS and O in ways that are not possible for Z, which is invariant and does not satisfy economy conditions of type (2iii) in any obvious way. It is somewhat paradoxical, perhaps, that logically equivalent versions of *t,* decomposing it in different ways, have important consequences. A further asymmetry is that * introduces only elements that are in [F] (though typically not in LEX), while the new elements introduced by Z never enter NS and are accordingly not in [F]; and if inclusiveness holds for NS, it introduces no features, even of [F].

[13] A still worse case is the EST model, with two cycles in NS, overt and covert.

[14] For a very strong version of the thesis, see Epstein 1999. On cyclic Spell-Out see *MI, DbP;* and for related ideas, Uriagereka 1999. S-O removes from NS all features that do not reach SEM. For expository simplicity, we refer to all these as "phonological."

[15] For an intriguing generalization of the notion to incorporate binding theory properties, see Freidin and Vergnaud 2001. For another possible generalization, see note 51.

[16] Possibly DP as well, but this raises many questions. I will put the topic aside here, keeping to the basic clausal architecture and the phases CP, *vV.*

[17] This requirement could reduce to Spell-Out of sister of the head if we adopt some variant of Ross's phonologically-empty performative analysis; Nissenbaum, pc.

[18] *MI* (21), DbP(ll). See also Nissenbaum 2000.

If ZP = [C [T *vP]*\, then T can access Quirky NOM object within *vP* (modifying its feature structure and also that of T),[1'] but C can access only the edge of *vP,* so that movement from the domain of *v* must pass through the escape hatch at the edge of *v*. PIC sharply restricts search and memory for $, and thus plausibly falls within the range of principled explanation (2ii, iii). It could be that PIC extends to NS as well, restricting search in computation to the next lower phase.

Let us focus attention on NS, taken to be the "generative engine" of L. Given LA, NS constructs a derivation D. LA is a set of elements of LEX. In the best case, these are "atoms" for D, undergoing no internal tampering in NS. Let us assume so, understanding this to mean that there is no feature movement and hence no "modified lexical items" (MLJs) with features attached to them.[20] That improvement is of some importance: feature movement is a complex operation, requiring some notion of "feature occurrence" that is not very clear; MUs also introduce many complications, best avoided if possible. Informally, we can refer to these atoms as the "heads of constructions."

NS has one operation that comes "free," in that it is required in some form for any recursive system: the operation Merge, which takes two elements *a,* P already constructed and creates a new one consisting of the two; in the simplest case, {a, P}.[21] The operation yields the relation € of membership, and assuming iterability, the relations dominate (contain) and term-of. The derived relation c-command (= sister of contain) functions at SEM (e.g., for binding theory), but perhaps not within NS. Any operation other than Merge requires empirical motivation, and is a prima fade departure from SMT (= (3)).

Informally, the new unit (a, P) is regarded as a "projection" of some head of a or p. In phrase structure grammars, including X-bar theories, the projection is identified by a new element (N, N-bar, NP, etc.), violating the inclusiveness condition. We therefore assume that (a, p} is identified either by a or by p (its *label);* a label, then, is always a head. In the worst case, the label is determined by explicit rule for each choice of a, p. A preferable result is that the label is predictable by general rule. A still more attractive outcome is that L requires no labels at all.[22]

If computation keeps to these austere conditions, it cannot rely on a head-to-SPEC relation R(H, SPEC); the relation called "m-command" in earlier work. There is no such relation. There is a relation R(SPEC, H), namely c-command; but no relation R(LB, H) where LB is the label of SPEC, since H is not in the minimal search space for LB (unless LB = SPEC).

[19] Here and below, assume Quirky Case to be inherent Case with a structural Case feature. Note that the NOM object neither raises nor undergoes any phonetic change in situ, the structural Case feature being invisible in this case. If properties of FL conspire to ensure that this is true generally of the domain of H in (6), the sister of H can be spelled out at HP (as in Nissenbaum's modification of PIC; see preceding note), with a universally invisible unvalued feature understood to be, in effect, a morphological convention rather than an actual feature, so that there is no crash at PHON. For expository convenience, let's assume that PIC can be formulated in terms of spell-out of domain (sister) of the head, ignoring these problems.

[20] Contrary to *MI* and earlier work, but not *DbP.*

[21] On binary vs. n-ary sets, see below.

[22] On the latter two possibilities, respectively, see *MI,* Collins 2001.

If operations are "driven" by labels, as we expect, then there can be no general SPEC-head relation at all, a strong and highly controversial conclusion to which we return. Computation driven by the label LB will keep to its domain, the category that guarantees minimal search, in accord with SMT.

In standard terminology, the first element merged to a head is its complement, later ones its specifiers (SPEC). In the best case, there should be no further restrictions on Merge; in particular, no stipulation on the number of SPECs, as in X-bar theories. There are further reasons to be skeptical about such stipulations. Typically, they are redundant; the limitations on Merge follow from selectional and other conditions that are independent. If empirical arguments are offered in support of restrictions on Merge, one must be careful to ensure that they do not follow from these independent considerations. It is sometimes supposed that stipulated restrictions have a conceptual advantage in that they reduce the number of possible configurations, but that is a dubious argument. Suppose we have a head H and three elements K, L, M to be successively merged to it. Free Merge yields the syntactic object SOi = {M, {L, (H, K}}} (L, M are SPECs of H). Stipulation that Merge can apply only twice yields SO2 = {M, {H', {L, {H, K)}}} (L the SPEC of H, M the SPEC of a new head HO- Stipulation that Merge can apply only once yields SO3 = {M, {H", {L, {H', {H, K})}}}, with two new heads H' and H". Each more restrictive stipulation reduces the types of possible configurations (under some interpretations), but there is no clear sense in which requiring SO3 is preferable to SO2, or either is preferable to SOi; if anything, the opposite would appear to be the case. Empirical arguments might be offered to show that H' and H" really exist, but if so, no restriction of Multiple Merge is necessary.

I will assume that there are no stipulated restrictions on Merge, and no projections or other violations of inclusiveness, keeping to "bare phrase structure."

Elementary considerations of efficient computation require that Merge of a to p ^ involve minimal search of P to determine where a is introduced, and least tampering with P: search therefore satisfies some locality condition (let us say, defined by least embedding; "closest" under c-command), and Merge satisfies an *extension condition,* with zero search. One possibility is that P is completely unchanged (the strong extension condition); another natural possibility is that a is as close as possible to the head that is the label of P, so that any SPEC of P now becomes a higher SPEC ("tucking in", in Norvin Richards's sense). Further questions arise under Merge with multiple SPECs.[24] Assume some version of the extension condition to hold, in accord with SMT.

The SM system requires that PHON indicate (ultimately temporal) order. A fairly standard assumption today, though not in earlier work, is that SEM involves only hierarchy, not order.[25] This version of IC is reasonable: let us adopt it — noting, however, that it is by no means easy to satisfy and is often violated in practice, even when adopted as a general principle. Holmberg's (to my mind, very convincing) theory of Object-Shift, for example,

[23] Note that the asymmetry is for expository convenience only; a and P are merged with no asymmetry.

[24] On these topics, see, inter alia, Richards 1997, BoSkovid 1998a, b, Pesetsky 2000, Nissenbaum 2000, and Brody (1995), on which much of this work relies, with reinterpretation. On tampering, see *MI.*

[25] On arguments to the contrary, see Fukui 2001.

requires explicit reference to the left border of a category (see Holmberg 1999, *DbP),* and other current work within a framework close to the one adopted here also requires left-right distinctions. The same is true much more broadly, but on the most austere assumptions about IC, such devices must be recast in different terms.

The C-I system requires that SEM express a variety of semantic properties. These include at least argument structure; call such properties "theta-theoretic," without commitment to one or another version of interpretability at the C-I interface. But beyond theta theory, C-I makes use of other kinds of semantic information, including scopal and discourse-related properties (new/old information, specificity, etc.)- The NS derivation therefore has to provide the basis for assignment of order at the SM interface, and for multiplicity of semantic properties at the C-I interface; presumably, one aspect of the diversity required by IC at SEM. Let us consider how these requirements are met.

Begin with order, determined by <I>. The worst case is that it is construction-specific. A better possibility is mat it is fixed once and for all for L: the head-parameter, along with a principle that determines that specifier (SPEC) precedes head —perhaps, as has sometimes been proposed, a reflection of a more general property that holds at other levels too (specifically, syllable structure: C-VC rather than CV-C), and may reduce to more general cognitive principles.[26] An alternative, developed by Kayne (1994) and a great deal of subsequent work, is that order reflects hierarchy. That approach eliminates the head-parameter, but at the cost of introducing many others (options for movement required to yield the proper hierarchies), and also some technical complications.[27] Hence the proposal requires empirical rather than conceptual argument, and that is the approach that has properly been adopted in pursuing these ideas. If correct, it appears to be a departure from SMT, contrary to what has commonly been assumed (by me in particular).

Let us turn to the multiplicity of semantic properties required at the C-I interface. Of these, the most fundamental are theta-theoretic properties (also incorporated in some fashion in artificial symbolic systems). Let us then reduce the multiplicity to duality: argument structure, and everything else. IC therefore imposes order at PHON and duality of semantic interpretation at SEM, with no interaction between <I>-PHON and Z-SEM.[28]

NS is based on the free operation Merge. SMT entails that Merge of a, P is unconstrained, therefore either *external* or *internal.* Under external Merge, a and p are separate objects; under internal Merge, one is part of the other, and Merge yields the property of "displacement," which is ubiquitous in language and must be captured in some manner in any theory. It is hard to think of a simpler approach than allowing internal Merge (a "grammatical transformation"), an operation that is freely available. Accordingly, displacement is not an "imperfection" of language; its absence would be an imperfection.[29]

» See Gil 1987, Carstairs-McCarthy 1999.
27 See Collins 2001, *MI,* Fukui 2001.
28 Another questionable assumption; see *MI* and sources cited.
29 Contrary to what I have assumed in earlier work. For over 40 years, there have been efforts to motivate displacement. That seems to have been a mistake. Recourse to any device to account for the displacement phenomena also is mistaken, unless it is independently motivated (as is internal Merge). If this is correct, then the radically simplified form of transformational grammar that has

The extension condition requires that displacement from within a be to the edge of a, yielding a new SPEC.

Internal Merge leaves a "copy" in place. Hence reconstruction is not an operation: it applies obligatorily in the base position. In A-movement, one position is selected for binding and scopal purposes; there is no need for a countercyclic operation of quantifier-lowering OL ^{x} Tnte "^{C0}Py theory of movement" is sometimes regarded as a controversial innovation. It is not: it is the null hypothesis.[31]

By definition, the operation TRANSFER (see (4)) applies at the phase level. At this level, internal Merge can apply either before or after TRANSFER, hence before or after Spell-Out S-O. ^^{e} former case yields overt movement, the latter case covert movement, with the *displaced* element spelled out in-situ.[32]

Covert and overt movement yield pairs <a, P>, a an edge element c-commanding P, where either a or P loses its phonological features under S-O: a under covert Move, P under overt Move. Understanding "copy" to cover both cases,

(7) K is a copy of L if K and L are identical except that K lacks the phonological features of L[33]

Both external and internal Merge are constrained in how they apply. We would like to show that the constraints are principled, deriving from (2ii, iii), not (i). It is unlikely that they have to do with PHON, which lacks relevant structure, so presumably they are imposed at the C-I interface, as conditions on SEM. There are two kinds of Merge (external and internal) and two kinds of semantic conditions at C-I (the duality noted earlier). We therefore expect them to correlate. That appears to be true. Argument structure is associated with external Merge (base structure); everything else with internal Merge (derived structure).

While plausibly regarded as the optimal outcome, the correlation is of course not a logical necessity. There could in principle be internal Merge to theta-positions, [M] and other devices might be employed to indicate scope and discourse-related properties: say, extra features on heads. But such devices have no independent motivation, and would also

become familiar ("Move-a" and its variants) is a kind of conceptual necessity, given the undeniable existence of the displacement phenomena.
30 Many questions arise about when and where reconstruction takes place. A particularly important contribution is Fox 2000. For some alternative conceptions, see Lasnik 1999c, Lebeaux 1999. For A'-movement, see particularly Bares 1986. Note that A- and A-movement have no status in the present framework; the terms are used only for convenience. It follows that no principles can be formulated in terms of the A-A'-distinction, a strong and highly controversial conclusion.
31 On the confusion of logic and history in this case, see *MI.*
32 ggg Njssenbaum 2000. See also Pesetsky 2000. I will assume this to be correct, contrary to *MI* and *DbP,* which interpreted covert movement as long-distance agreement.
33 o^ -phonological" as used here, see note 14. Since we are assuming a "classical LEX," with phonological and semantic features of roots included, pure synonyms are distinguished in the lexical array (numeration), so no false chains will be introduced.
34 See, e.g., Horstein 1999.

require new rules. In contrast, internal Merge is free, so on minimalist assumptions it is to be expected that FL will use this device, as appears to be the case. FL takes scopal and discourse-related properties to be "edge phenomena" (hence involving c-command).

Uncontroversially, theta-theoretic properties depend in part on configuration and the semantic properties SEM(H) of the head (label). In the best case, they depend on nothing else (the Hale-Keyser version of theta theory). Assuming so, there are no s-selectional features or theta-grids distinct from SEM(H), which is typically a rich and complex structure,[35] and theta-theoretic failures at the interface do not cause the derivation to crash; such structures yield "deviant" interpretations of a great many kinds. The only other possibility is merger of a semantically-uninterpretable element in a configuration that lacks a theta-theoretic interpretation, in which case it will have to be deleted before SEM. Reference to s-selection below is to be understood as informal reference to the effects of SEM(H) in particular configurations.

Elimination of s-selection and related notions is motivated by other considerations beyond their redundancy in pure-configurational theta theory. Suppose that the transitivity marker *v* has an s-selection feature F requiring an external argument EA. One immediate problem is that the requirement is formulated as a head-to-SPEC relation, and we have seen good reason to believe that this cannot exist (nor the broader symmetric SPEC-head relation, in the general case). But even extending relations to m-command leaves problems. Suppose we have derived (8):

(8) *v* [see [the picture]]

To satisfy F, EA must be merged before the derivation moves on, or F will remain unchecked (because of cyclicity of Merge) and the derivation will crash at SEM. But this condition is unnecessary, therefore unwanted: it amounts to stipulating a part of Burzio's generalization, which follows independently from Case-theoretic considerations.[3]* These problems become more severe if there are no categories, only roots, so that *v* in (8) determines that the root *see* is verbal —on many grounds a reasonable assumption, which also yields the otherwise unexplained conclusion that V -> *v* movement is obligatory. Then given, say, the root *arrive,* we do not know whether it is verbal (selecting an internal argument IA) or nominal (with no IA required) until the next stage of derivation, at which point it is too late to merge IA (by cyclicity). There are still further problems; e.g., how do we know which s-selectional feature must be satisfied first?[37] For a variety of reasons, then, s-selection should be dispensable.

35 There is nothing to say, then, about the fact that they are not morphologically manifested, though argument position may be marked morphologically, a different matter.

36 If the internal argument receives Case, a higher element that requires agreement (T or *v),* which must appear in a well-formed construction for independent reasons, will not be able to have its features checked. We assume that once Case of a is checked, a is "frozen"; it cannot enter into further agreement relations. See below.

37 See Collins 2001, who also proposes that c-selection is a head-head property that avoids labels, with some interesting empirical consequences, but some problems as well. Note that nothing requires that EA be an actual argument: it could be a *there-type* expletive with <(>-features (similar to French il). For a possible example from Greek, see Iatridou 2001.

Elimination of s-selection has a number of consequences: it undermines at least the simplest ways of predicting labels (as in *Ml),* requiring a restatement in terms of SEM(H). It also entails that derivations cannot be failure-proof ("crash-free"), thus undermining the strong version of IC mentioned earlier. But these cannot be reasons for insisting on s-selection features, plainly. We have to find other ways to determine labels, or show that they are dispensable. And IC must be weakened. Perhaps the condition should be that L yields an infinite variety of prepositional configurations (CP or *vP),* and that each element of LEX enters into these.

Properties of the C-I interface, then, determine generally the application of external Merge. If we dispense with s-selectional features, failure to satisfy the selectional properties they were taken to express does not block convergence but yields deviance, an outcome that can be detected instantly in the cyclic NS derivation.

What about internal Merge? We expect its application to be motivated by the nontheta-theoretic C-I conditions: scopal and discourse-related (informational) properties in particular. That appears to be the case as well. Scope has the familiar "long-distance" property: scope of *wh-,* for example, can be well outside its phase. Given PIC, it follows that internal Merge (movement) must be successive-cyclic, passing through the edge of successive phases. The same is true of discourse-related properties. Note that these properties are not built into artificial symbolic systems, which need not satisfy IC and do not resort to internal Merge.

The extra edge position in a required by internal Merge is optional, and has no theta-role. Assuming options to be determined in LEX, the head H of a must have a feature that makes this position available: an EPP-feature in standard terminology; from another point of view, the feature OCC that means "I must be an occurrence of some P".[38] Optimally, OCC should be available only when necessary: that is, when it contributes to an outcome at SEM that is not otherwise expressible, the basic Fox-Reinhart intuition about optionality.[39] Hence H has OCC only if that yields new scopal or discourse-related properties (or if required for other reasons; see *DbP).* No non-local or look-ahead conditions are introduced. If H has OCC, then the new interpretive options are established if OCC is checked by internal Move; it is only necessary that the cyclic derivation D can continue so that they are ultimately satisfied with convergence of D. Informally, we can think of OCC as having the "function" of providing new interpretations; in the analysis of any process or action (the operation of the kidney, organizing motor action, generating expressions, etc.) such functional accounts are eliminated in terms of mechanisms.[40]

Note that no SPEC-head relation is involved. The new interpretive options result from checking of OCC by internal rather than external Merge, the former a reflex of a head-head relation in the *MI-DbP* framework assumed here; it is not the SPEC-head relation, but the way it is satisfied that is crucial. It therefore conforms to the empirical thesis (9):

38 Taking an occurrence of P to be its sister, as in the theory of chains without indices; see DbP and sources cited.

39 *Ibid.*

40 *Ibid.* See also Lasnik 1999b, Frampton and Guhnann 2000.

(9) Apparent SPEC-H relations are in reality head-head relations involving minimal search (local c-command)

The thesis conforms to SMT, but faces serious empirical challenge.

IC requires that all features be interpretable. But it is clear that there are uninterpretable features that must somehow be eliminated before the NS derivation is transferred to *Z,* a prima facie "imperfection" that appears to fall within the unexplained category (2i). One example is the OCC feature, but insofar as it can be accounted for in the terms just outlined, it moves to the preferred category (ii)-(iii).

Case-agreement systems involve a richer array of uninterpretable features, and are of particular interest for this reason in the minimalist context: structural Case for nouns, <p-features for categories that agree with nouns; assume these to include T for subject-agreement and *v* for object agreement. "These must be eliminated in NS. Let's first consider some of their general properties, then ask why they exist.

Uninterpretable features are eliminated when they satisfy certain structural conditions: an uninterpretable feature of *a* must be in an appropriate relation to interpretable features of some p. Furthermore, P must be *complete,* with a full set of features. Nouns are always complete, since their <p-features are always present (and interpretable); hence nouns check the <p-features of agreeing categories. Participles are not complete (lacking person), and do not check Case.[42] T may be complete or *defective;* if the latter, it does not check Case. Feature-checking, then, resolves to pairs of heads <H, H'>, where at least one is complete and they are in an appropriate relation. For optimal computation, one member of the pair must be available with no search. It must therefore be the head H of the construction a under consideration, a = {H, XP}. Call H a *probe* P, which seeks a *goal G* within XP; P = H c-commands G, but that is a consequence of minimal search. If the P-G relation satisfies relevant conditions, then uninterpretable features of P, G delete.

We therefore conclude that in addition to Merge, there is a relation Agree holding between probe P and goal G, which deletes uninterpretable features if P and G are appropriately related. It remains to determine its properties.

These considerations lend further support to the conclusion that the SPEC-head relation does not exist But there is strong empirical evidence that it does exist. Much work

[41] With provisos discussed in *DbP* and elsewhere. We keep here to Nominative-Accusative systems, but the same reasoning should apply to Ergative-Absolutive systems. Theta-related (inherent) Case is a separate matter.

[42] **But a participle PRT may have Case, shared with that of its object OB. See *DbP* for discussion within a phase-theoretic framework, with Case** [illegible] **T, ACC with** [illegible] **If OB has only** [illegible] **. UI.IEIUUIKU *vy* me mgner i^ase-assigner (NOM wit]** ,.----------->w.u/ TMuuui<uv.a5e, the PRT-OB relation in itself cannot determine the (uninterpretable) Case of either, but if OB has inherent Case, the relation suffices to determine case of PRT, as in the Icelandic counterpart of (i) (from Boeckx 2000):

(i) me(DAT> seemS(Defouit) t[John(DAT) bebelieved<NOM,pi) like horseS(NOM,pu

Here OB (= *horses)* has inherent Case, optionally Quirky (but not here).

relies on SPEC-head relations to provide positions for surface phrase structure.[43] It is also well-established that raising tends to yield richer visible morphological realization of inflection than long-distance agreement.[44] We therefore either conclude that the unexplained category (2i) is more comprehensive than we would like, or we have to reanalyze this evidence in other terms. In some cases, this may be fairly straightforward. For example, morphological richness of H in the SPEC-H relation (as compared with long-distance agreement) is not actually a reflex of the SPEC-H relation but of the way it is satisfied: by internal rather than external Merge. It therefore correlates with the new interpretive options provided by H when OCC (= EPP) is satisfied by internal Merge; in both cases, we are dealing with the same property of H, and the SPEC-head relation plays no role, conforming again to the thesis (9). A similar observation holds for successive-cyclic A'-movement that yields a special form of C at the intermediate and final positions, as in Irish. McCloskey (2000c) shows that the forms of C are determined not by the SPEC-head relation but by the way it is satisfied: external or internal Merge, again conforming to (9).[45] The more general task is not trivial, and remains an interesting research project.

If mere is no SPEC-head relation, then the EPP-feature OCC cannot be satisfied by Merge alone. It follows that internal Merge requires Agree. Therefore, Move = Agree + Pied-piping + Merge. Note the weakness of the hypothesis. It would be refuted only by a configuration H-XP in which any arbitrary term of XP could raise to SPEC-H. But it seems that raising of a from XP is always restricted to some category of constituents of XP, hence some feature F of a (or complex of features) that matches OCC. The (nontrivial) question then reduces to what F is. «*

It also follows that external Merge does not suffice to check OCC. The only relevant case is expletive EXPL. EXPL externally merged in SPEC-T must delete the OCC feature of T and lose its own uninterpretable features (if T is complete). The interesting case is a *there-type* EXPL lacking theta-role. EXPL must have some feature [uF], or it could not be raised.[47] Suppose EXPL is a simple head, not formed by Merge. In a label-free system, EXPL is accessible without search as a probe, and can match and agree with the goal T. If T is selected by C (hence complete), then [uF] is valued and disappears, and the derivation can converge. If T is defective, EXPL will await a higher complete probe (either C-T or *v).* Whatever probe values and deletes [uF] must still seek a complete goal, to eliminate its own

[43] One important recent discussion is McCloskey 2000b.

[44] See Guasti and Rizzi 1999. For extensive evidence that SPEC-head agreement not only exists but plays a decisive role, see Chung 1998.

[45] McCloskey also provides evidence that the resumptive pronoun strategy involves external Merge of pro in SPEC-C, yielding an operator-variable structure that may form "mixed chains" with raising. Hence the attempt in *MI* to describe (not account for) appearance of EXPL in SPEC-T but not SPEC-D in phase-theoretic terms cannot be correct. The problem of accounting for the distribution of EXPL in some principled way (noted some years ago by Julie Legate) remains open.

[46] One might seek to appeal to universal conditions C: *a* is allowed to raise only if it satisfies C. Even if this is possible, to show that Merge alone checks OCC it would be necessary to show that C does not invoke head-head relations (as in standard formulations of MLC). A technical question is how checking of the EPP-feature by internal Merge (Move) is effected by a head-head relation alone. Neither Agree nor (by assumption) Merge can check the feature, so it must be a property of Pied-piping, still in many ways a mysterious operation.

[47] Perhaps, for example, structural Case, as proposed by Lasnik 1999a, chap. 4.

uninterpretable features: the normal case of Mere-constructions with long-distance agreement. Suppose that EXPL has all (p-features, like French *il.* When merged in SPEC-T, T complete, it can no longer raise: therefore T must value and eliminate the (p-features of EXPL. But that can happen only if T finds a goal to value its own features - which, however, are overridden by the EXPL-T relation (possibly a reflection of the property of richness of morphological realization already mentioned). We conclude that such expletives must be simple heads, and that there is an additional empirical argument in favor of Collins-style label-free phrase structure —noting that some problems remain unresolved at least in any clean way.

If the head H of a has the feature OCC, then something must be Merged in SPEC-H to check and eliminate it. If external Merge of EXPL is inapplicable, internal Merge must find an agreeing active goal G, which induces Pied-Piping to yield K(G), then merged in SPEC-H. Suppose G is unable to check and delete the features of H; e.g.. Quirky Dative G that raises to SPEC-T but does not satisfy the requirement that the features of T can be checked only by Nominative. Then T can either have a default inflection if the language allows that option, or can find a lower active Nominative with which it can agree, deleting its uninterpretable features.[48]

Covert movement to the escape hatch SPEC-PP is possible for a direct object only if it undergoes further A'-movement (in the informal sense: see note 30). Thus there is covert to/i-movement, but not covert Object-Shift OS (yielding the semantic edge properties but without overt movement). If OS is Case-driven, and Move includes Agree, then we cannot have the sequence of operations: Agree(p, Object), TRANSFER, OS. That would require that Agree apply both before and after TRANSFER (specifically, S-O).[49] But utfi-movement is plainly driven by a different feature, as successive-cyclic and adjunct movement make clear. Therefore it can apply (covertly) in a unitary fashion after TRANSFER.

In the probe-goal system, or any approach based on Attract rather than Move, it follows from optimal computational considerations that Merge must be binary, minimizing search for the goal. The conclusion has been generally assumed, but has resisted explanation and is not obvious; some considerations might yield a preference for n-ary categories.[50]

Returning to the relation Agree, the problem is to show, if possible, that it is reducible to categories (2ii, iii), thus conforming to SMT.

The simplest version of Agree would be based on the free relation Match: identity of features. For Agree to delete the features of P or G, the paired element (G or P, respectively)

48 See *DbP* and sources cited for further detail. For much more extensive discussion of related issues, see Boeckx 2000.

49 Possibly this falls together with the Maximization principle (see *DbP,* (14), (21)). Recall that Agree must apply to the direct object before S-O because of phonetic reflexes and to remove the uninterpretable Case feature from NS, and that Case-driven movement covers Quirky Case as well. On problems concerning A'-movement, see note 51.

50 That was assumed in early efforts to motivate transformations in processing terms; see Miller and Chomsky 1963 for review. On binary structures, see particularly Kayne 1981.

must be complete. Furthermore, P and G must be *active:* once their features are checked and deleted, these elements can no longer enter into the Agree relation; the Case-checked subject of a finite clause, for example, cannot check uninterpretable features of the next higher phase head or raise to this position. To minimize search, the P, G relation must be local. In *DbP,* it is assumed that G must be the closest matching H, but there is good reason to believe that like others, this property must be relativized to phases, so that P can find any matching G in the phase PH that it heads, simultaneously deleting uninterpretable features. It follows that intervention effects will hold only if the intervening element is not rendered inactive by P itself.[51]

This modification is both natural and empirically motivated. It means that the operations driven by the head H of a - Merge (external and internal) and Agree — are in effect simultaneous. We have already implicitly made this assumption by taking the probe H in a = (H, XP} to be detectable with no search, a might of course have SPECS, but these are added simultaneously with probe for G.

It is unexpected that *v* and T (rather than the two phase heads *v* and C) should be the probes for the Case-agreement system. The departure from expectation is an illusion, however. T functions in the Case-agreement system only if it is selected by C, in which case it is also complete.[52] Furthermore, in just this case T has the semantic properties of true Tense. These cannot be added by the cofeatures, which are uninterpretable; they must therefore be added by C[53] Hence T enters into feature-checking only in the C-T configuration, and the symmetry is restored: the two phase heads C and *v* are the operative elements. There is further evidence to support this conclusion. Successive-cyclic A'-movement often leaves a reflex, sometimes in C (where we would expect it), but commonly in the agreement system headed by T (where we would not).[54] That makes sense if C-T are really functioning as a unit in inducing agreement.

We would expect to find similar properties for *v.* A possible example is Indonesian. There is a transitivity marker and two options for irfz-questions: in-situ or successive-cyclic

51 Frampton et al., 1999; Hiraiwa 2000. See the latter for extensive discussion of intervention effects and multiple Agree, with parametric variation. The former points out that this proposal can overcome a problem they note in the analysis of participial constructions in *DbP,* which works if the object remains in situ but fails if it raises, blocking Case-checking of the participle; as they note, this could also be overcome along lines of the proposal at the end of *DbP,* which dispenses with literal movement. Multiple-Agree eliminates unvalued features of traces, as desired. While this is the right result for A-movement, it may not be for A'-movement; e.g., the utfi-feature of a trace is not valued until a higher phase. The asymmetry reflects another one: all A-movement properties are handled within a phase, but not A'-movement properties. That suggests that there might be a more abstract notion of phase, based on the concept of valuation of features rather than just size of the category, as here and in earlier work.

52 There are unstated assumptions here, e.g., that control structures are CPs.

53 Additional evidence is provided by languages like Greek, in which T with a complete set of co features but not selected by C lacks semantic tense and is defective with regard to Case and agreement; Iatridou 1993 (1988), and much subsequent work.

54 See Chung 1998 for review; also Collins 1993. Irish is the best-studied case of C-reflex. See McCloskey 2000c. Inflectional marking of C in some languages yields further support. For more on C-T relations, assuming literal head movement, see Pesetsky and Torrego 2001.

movement. With *wh-* in-situ, the marker remains, with the movement strategy it disappears at each stage.[8]

An uninterpretable feature F must be distinguished somehow in LEX from interpretable features. The simplest way, introducing no new devices, is to enter F without value: e.g., [uNumber]. That is particularly natural because the value is redundant, determined by Agree. Therefore Match is non-distinctness rather than identity. Uninterpretable [uF] has several important properties: (1) it must be valued under Agree for the NS derivation D to converge; (2) once valued, it must be eliminated from D; (3) it must be transferred to <I> by TRANSFER before it is eliminated, since it may have a phonetic reflex. Furthermore, (2) must be carried out quickly, without need for search to earlier stages of derivation. The optimal way to deal with all properties is to regard valuation of [uF] as, in effect, part of TRANSFER. This operation removes features that would cause D to crash at SEM, including phonological/morphological features of LEX, and [uF] if it has just been valued. We therefore conclude that TRANSFER (hence S-O) must be cyclic, confirming the earlier conclusion based on computational efficiency. Note that TRANSFER has a "memory" of phase-length, meaning again that operations at the phase level are in effect simultaneous. It follows that phases should be as small as possible, to minimize memory for S-O, and independently, to maximize the effect of cyclic derivation in simplifying O.

A major problem is why uninterpretable features and Agree exist at all. We have an answer for OCC at the phase level, but not for OCC at T: the original Extended Projection Principle — perhaps universal, perhaps not; the jury is still out on that, I think.[56] More important are the features of the Case-agreement system. These features fall into three types:

(10) (i) cp-features on the probe (T, *v)* (with T tensed and complete, i.e., really C-T) (ii) structural Case on the goal (N or D) (iii)EPP-feature (OCC) on the probe.

A suggestive fact is that internal Merge requires just these three kinds of information. The target is determined by the probe (i), which also determines what kind of category can be merged to it. The moved element is determined by the goal (ii) (which has to be active). The availability of a position for Merge is determined by (iii) (which may allow long-distance agreement if the feature is valued in some cheaper way, and may allow multiple

[55] Aguero 2001, data from Saddy 1991. A host of further questions arise in this connection, including differences of richness of inflection of T, *v,* and correlated differences of overt Case-marking, in Nominative-Accusative vs. Ergative-Absolutive languages.

[56] For English it appears to hold invariably for T, as we can see in raising constructions (awkward, because there are so few relevant raising verbs, but the point is clear), as in (i), *U* the trace of *John:*

(i) John seems toY [ti to appear toX [t2 to like Mary]]

Suppose Y = *her.* At the relevant level (presumably SEM), Y c-commands *Mary,* inducing a Condition (C) effect. It therefore c-commands X, which can however be *her* linked to Y, or *himself* linked to *John* (despite intervening Y); both facts imply that fi = *John.* Therefore intermediate T, though defective, must satisfy EPP. For argument that Irish does not have EPP, see McCloskey 2000b.

subject if the language allows OCC to be satisfied in both of the possible ways, by external and internal Merge, as in Icelandic).

All of this falls into place if uninterpretable features are the mechanism for displacement, perhaps even an optimal mechanism. But displacement comes "free," and its application is determined by IC: the duality of semantic interpretation at SEM. If this line of reasoning is tenable, then uninterpretable features and the extra relation Agree move from the unexplained category (2i) to the category of principled explanation (ii-iii), and the discussion so far continues to conform pretty closely to the strong minimalist thesis (3).

Cyclicity of derivation requires that Merge to a always be at the edge of a, satisfying an extension condition, strong or weak ("tucking in"). Non-cyclic Merge to a term properly contained in a complicates all three parallel derivations: NS, G>, Z. There appears to be one significant counter-example to cyclic Merge: late insertion of adjuncts,[57] as in (11):

(11) [wh- which [$_o$ [NP picture [p of Bill]] [ADJ that John liked]]] did he buy t_{Wh}-

Despite complexities, the tendencies are fairly clear: linking of *Bill* to *he* induces a Condition (C) effect, but Unking of *John* to *he* does not. The effect for *(he. Bill)* is expected by (obligatory) reconstruction, but not its obviation for *(he, John).* That would follow, however, if adjuncts can be late-merged at the root, though not complements, consistent with the fact that the complement P is s-selected but the adjunct ADJ is not.

Is there a way to deal with these problems while preserving cyclicity in a single cycle in accord with SMT (= (3))?

More fundamental questions arise at this point. There has never, to my knowledge, been a really satisfactory theory of adjunction, and to construct one is no slight task. That has been recognized since the earliest efforts to construct a reasonable version of phrase-structure grammar, what came to be called X-bar theory.[58] 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]. The construction is crucially asymmetric: if a is adjoined to P, the construction behaves as if a isn't there apart from semantic interpretation, which is not that of standard X-bar-theoretic constructions; island properties differ as well. P retains all its properties, including its role in selection. There is no selectional relation between P and a, so determination of label evidently relies on the asymmetry to determine, say that "the picture of John's" is an DP, not a PP, while "the picture of John" has an N-PP complement; and "destruction of the army" is one or the other, in one case interpreted as in "the achievement of the army." The adjunct a has no theta-role in <a, P>, though the structure does - the same one as p.

[57] Proposed by Lebeaux (1988) to deal with problems noted by Robert Freidin, and Henk van Riemsdijk and Edwin Williams.

[58] See, e.g., the discussion of "house in the woods" vs. "book in the woods" and other such matters in Chomsky 1970.

Suppose we see how far we can go starting from first principles, assuming SMT, and keeping to core problems, concentrating on what should be captured by some eventual formalization.

For structure building, 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 P and a and forms the ordered pair <*a,* P>, a adjoined to p. 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 a as attached to P on a separate plane, with P retaining all its properties on the "primary plane," the simple structure.

Two questions arise about adjunction: (1) Why does it exist? (2) How does it work?

The natural place to seek an answer to question (1) is at the SEM interface. Recall that the strong interface condition (however formulated precisely) requires sufficient diversity at SEM. Possibly richness of expressive power requires an operation of predicate composition: that is not provided by set-Merge, which yields the duality of interpretation discussed earlier: argument structure and edge properties. But it is the essential semantic contribution of pair-Merge. If the C-I system imposes this condition, then the existence of a device to yield predicate composition would conform to SMT —a promissory note, given the limitations of understanding of C-I, but not unreasonable.

Let's turn to the second question: How does adjunction work? We have to determine how relations and operations apply to complex objects formed by asymmetric pair-Merge in such a way as to capture the basic properties of the construction, keeping to SMT as far as possible.

Assume that like other operations, adjunction of a to p applies cyclically. P behaves throughout as if it were in a simple structure formed by set-Merge. We can implement the observation by taking P to be in fact set-merged in the standard way; adjunction then applies to replace p by «x, P>, where the information that P is set-merged is captured by the asymmetry (it is part of the interpretation of the pair). The semantic role of <cc, P> is determined compositionally under Z. In (11), for example, ADJ is pair-merged to NP in the base position, and [DET <ADJ, NP>] receives its theta-role in the normal way, with composition of the predicates NP, ADJ. We take [DET <a, P>] to be "in a configuration" at SEM, but that seems unproblematic: "in a configuration" is not one of the relations defined for simple structures, and the assumption here is as natural as any. Any role played by P in selection or other semantic interpretation is preserved, because it is interpreted to be set-merged, in a simple structure.

What about Condition (C) at SEM? When X c-commands <cc, P>, does it also c-command *a* and P? P was introduced by set-Merge, and before a was adjoined to it, X c-commanded p. But the central property of adjunction is that adjunction of a to P does not change the properties of p. For P to lose some property when a adjoins to it would be a complication, an "imperfection." The relation c-command(X, P) is therefore not lost when

a is adjoined to P: accordingly, X still c-commands p in <a, P>, as before adjunction. But extension of c-command to the adjoined element a would be a new operation, to be avoided unless empirically motivated. Happily, the empirical evidence disconfinns the complication. Cases of the type (11) fall into place, insofar as they are dear.[59]

We know that at the stage where <a, P> is spelled out, it also becomes a simple structure at SEM. Thus if (11) is embedded in the context "he asked -," *John* is subject to Condition (C) and hence must be c-commanded by matrix *he.* Therefore there is an operation SIMPL that converts <oc, P> to {a, P}; in effect, it is part of Z. 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 (= (4)), which transfers the NS derivation (specifically, its last line) to both <t> and Z.[60]

Suppose SIMPL is optional. Recall that overt movement, as in (11), requires the ordering of operations: Move-TRANSFER. For overt movement, then, optionality of SIMPL will have no effect at the PHON level, because S-O does not apply to the trace in any event.[61] But it might have an effect at SEM. Thus in such structures as (11), 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. In the case of covert movement, with the ordering TRANSFER-Move, S-O applies and therefore SIMPL feeds Z as well.

The account so far is consistent with the basic assumption throughout that ordering is required only at PHON. Within O, a of «x, P> is integrated into the primary plane, in the informal version.[62] Along with S-O in general, ordering is not part of NS. It can therefore apply no earlier than TRANSFER. If it applies later than TRANSFER, computation will be more complex, since it will have to apply separately in the mappings <E> or Z. We therefore assume that SIMPL applies at the point of transfer, so that SIMPL — an optional component of TRANSFER - is in effect part of both O and Z. SIMPL converts <cc, P> to {a, p|, which is then ordered, using the information that this is pair-Merge, not set-Merge (unproblematic, because SIMPL is in effect part of S-O). The instantaneous and uniform operation Z retains the same information at SEM, so that even when simplified, <cc, P> is interpreted as adjunction.

It follows that in the structure <a, p>, a is integrated into the linearly ordered structure at the stage of derivation where P is spelled out. Perhaps a is adjacent to P, perhaps not; that depends on properties of adjuncts, which are not all the same in this regard. These are separate issues, which have to be addressed however adjunction is

[59] Judgments fade as the r-expression in the adjunct *(John* in (11)) is more deeply embedded. It is tempting to relate this to the fact that deeper search is required at the root to spell out the pair <a, (5>, along the lines to which we turn directly.

[60] We are assuming that S-O applies at the point of transfer: that is, there are no phonological operations that must apply before S-O (and ordering). If there are, the account becomes more complex, but unchanged in essentials.

[61] Apart from the possibility mentioned in note 64.

[62] Note that a may have already been spelled out, in part at least, prior to adjunction. The question is how and where it is integrated into the simple configuration that is spelled out. Recall that lots of computation is proceeding in parallel.

handled. The crucial point is that we will not find P spelled out in one place with a appearing somewhere else as a result of movement of P - either covert Move carrying a along or overt Move leaving it behind. In short, the principle (12) holds:

(12) In <*a,* P>, a is spelled out where P is

Suppose (12) is rejected: integration of the adjunct a can be dissociated from spell-out of p. That raises all sorts of conceptual problems, but even if these can be resolved, the proposal requires a new rule INSERT that applies to *«x,* Copy(P)>, inserting a into the simplified structure.[63] These complications are unwanted, and seem empirically incorrect.

Consider successive-cyclic movement, applying to an underlying expression of the form (13), with ADJ-P = "with great annoyance" or "every day of his life" (or both, to bring out the intended interpretation with ADJ-P understood as qualifying "remember"; the extra phrase "he heard" is inserted to avoid Mat-trace issues):

(13) Bill remembered [that he heard Qohn had insulted him]] ADJ-P

Now replace *John* by *which person,* which raises to the root either overtly or covertly, yielding (14), where *wh-i = Copy(which person)* if raising is overt, and *wh\ = Copy(which person)* if it is covert (in the latter case, take *BUI* to be replaced by *who,* raising to SPEC-C):

(14) *wh-i* (1) [Bill [(2) remembered [(3) that he [(4) heard *[wh-2* (5) [had [(6)
insulted him (7)]]]]]]] ADJ-P

Whether overtly or covertly, *which person* passes through positions (l)-(6).

Suppose REL = "who taught at Harvard" is adjoined to *which person,* necessarily in the base position (6). Suppose dissociation of Spell-Out is allowed (contrary to (12), the extra rule INSERT exists, and the conceptual problems are somehow overcome. Then REL should be able to appear at PHON at positions (l)-(6) - or (7), if INSERT places the adjunct to the right. But it can appear only at (1) (if *wh-\ - which person,* after overt movement) or at (5) (if *wh-i = which person,* with covert movement; possibly (7)). That is, REL must be spelled out where *which person* is spelled out; any other choice yields severe deviance. Therefore we can avoid the conceptual problems and the new rule INSERT: SIMPL is part of TRANSFER with no further complication, and Spell-Out accords with (12).

Let us turn to application of Move to <*a,* P>. Here a preliminary question arises: do we take Move to be literal internal Merge (overt and covert movement), or do we adopt the approach outlined at the end of *DbP,* with no internal Merge but only marking of the OCC feature of the probe? If the latter, no special problems seem to arise, but let us continue to adopt internal Merge. Clearly <a, P> can be moved as a unit, as in (11), where SIMPL applies at the root. That means that the operation of pied-piping, triggered by the active element of P (the w/i-feature in (11)), can pick up the adjunct along with everything else in the category it identifies. Exactly how pied-piping works remains somewhat unclear, with well-known problems still outstanding, but it must at least have this property. Is there also an option for pied-piping to keep just to the simple part of the construction, hence moving P while leaving a in situ? That would violate (12), yielding the dissociation of S-O that we have just rejected on conceptual and empirical grounds, so we dismiss this option.

Note that the same considerations apply to the countercyclic late-Merge approach to adjunction of a to P: it should be permitted only where P is spelled out, in accord with (12). Merge cannot apply to a copy: a trace or an empty category that has moved covertly.[64]

If this line of reasoning is correct, then the complications of countercyclic late Merge can be eliminated at least in the standard cases. Given the (presumed) interface condition that requires compositionality of predicates, the properties of these structures follow from minimal conditions on SIMPL: optional application at TRANSFER. Furthermore, (12) holds.

More generally, we have the basic outlines of a theory of adjunction that satisfies SMT.

The late-Merge operation for adjuncts has recently been exploited in work by Danny Fox and Jon Nissenbaum to deal with such constructions as (15), in which the adjunct is apparently extraposed in conflict with the general principle that only complements, not adjuncts, can be extracted from NP (we ignore NP-DP distinctions where irrelevant); and (in (ii) the Condition (C) effect is obviated for the pair *(he, John),* though it would be expected (by reconstruction) if ADJ is extraposed from NP:[65]

(15) (i) we saw [NP a painting] yesterday [ADJ from the museum]

(ii) I gave him [NP a painting] yesterday [ADJ from John's collection]

Fox and Nissenbaum argue that strange properties of these and many more complex constructions can be explained if *a painting* is covertly raised to the right (by QR), and ADJ is then merged post-cyclically.

Though the results are impressive for a wide range of constructions, there are a number of problems. One is that late-Merge is employed. Possibly the analysis can be reconstructed in terms of cyclic adjunction, but even if so, other problems remain. Dissociation of Spell-Out of adjunct and host is required in violation of (12), but that is

[63] Copy is asymmetrical: Copy(P) is p stripped of its phonological features: either the trace of p or covertly-raised p. See (7). Since operations at the phase level are in effect simultaneous, the asymmetry of Copy is determined by TRANSFER, Move, and their ordering at the phase level.

[64] Perhaps Merge cannot apply to any empty category ec, not really a generalization if a copy is just a syntactic object with no phonological features. Alternatively it could be that trace is destressed, entering into ellipsis under the usual paradigm, but obligatorily in this case. Here questions arise about apparent phonetic effects of trace vs. lexically-determined ec's. Barring of adjunction to ec's seems correct for PRO and pro, but there might be other reasons. Note that phonological features do not affect NS-derivations, but their presence or absence may (see note 33). On this matter, see *DbP.* The conclusion would leave intact the results of Distributed Morphology, while abandoning the condition of obligatory late insertion for roots, questionable on other grounds; see *DbP.*

[65] See Fox and Nissenbaum 1999, Fox 2001, and Nissenbaum 2000 for extensive discussions and references to earlier work of Peter Culicover-Michael Rochement and Tarald Taraldsen.

problematic, as just discussed. It is also unclear why QR is to the right; a covert operation should have no ordering properties. One cannot - at least in any obvious way — appeal to the fact that authentic extraposition is to the right, since a crucial (and well-motivated) part of the analysis is that there is no adjunct extraposition; nor can Heavy-NP-Shift (HNPS) be invoked, because of the dissociation of Spell-Out of *a painting,* not true for HNPS. There is also a conceptual question: apart from serving as an empty bearer of adjunction, QR typically does not feed G>. It should, then, not be part of NS, just as ordering is not.

An alternative approach is suggested by the fact that very similar expressions are generated independently, namely, those that introduce qualifications or afterthoughts, as in)

(16) (i) we saw [NP a painting] yesterday, (that is,) a painting (one) [ADJ from the museum]

(ii) I gave him [NP a painting] yesterday, (more precisely,) a painting (one) [ADJ from John's collection]

Here "a painting" is destressed in the adjoined phrase and can undergo ellipsis in the normal way, yielding (15). The scopal and other properties of (15) follow without recourse to QR or countercydic Merge. There is no need to violate the theoretically well-motivated and apparently empirically valid principle (12). It also follows that the adjunct is to the right. The adjunct-complement distinctions that Fox-Nissenbaum describe are also accommodated: if the "afterthought" is, say, "a [picture of his father]," then "a picture," not being a unit, cannot delete.

Fox applies similar ideas to the intricate problems of antecedent-contained deletion (ACD). Still keeping to the simplest case, consider (17):

(17) John [VP likes [NP every boy Mary does <likest>]]

Here <...> is elided with VP as antecedent. But the parallelism requirement for ellipsis is violated. For such reasons, ACD-resolution is standardly assumed to involve QR, raising NP. But this solution is inconsistent with the simplest theory of movement (the "copy theory"), which reintroduces the problem. Furthermore, condition (C) effects would be induced if movement leaves a copy, but they are not.

Fox develops a very simple analysis along the lines of the analysis of (15) just reviewed, and shows that it deals with a wide range of cases. The base-generated form for (17) is (18):

(18) John [vp likes [NP every boy]]

[66] We know that these are to the right, independently (and for obvious reasons), because of the variety of such constructions; the simplest of them are relevant here, but the others clearly exist.

QR raises NP to the right, at which point the relative clause is late-merged. Parallelism is straightforward; there are no condition (C) or other reconstruction effects; scopal properties and complement-adjunct distinctions fall into place.

Though the results are again impressive, the same problems arise as for the (15). It is natural to ask, then, whether the alternative approach just outlined can overcome them. Suppose that the underlying structure for (18) is (19) — which, again, is generated independently as one of a large class of cases:

(19) John likes every boy, (that is, more accurately,...) every boy Mary likes

We derive essentially the same results without the problems associated with late-Merge, violation of (12), and QR. There might also be other advantages: ACD-resolution requires application of QR in constructions in which it is usually barred.[67] These problems might be obviated if the scope of ACD is determined in the manner suggested here, without QR. If something like this approach is tenable, then ACD essentially disappears as a phenomenon.

This discussion only scratches the surface.[68] A great many important questions remain, but it seems that cyclic Merge is at least compatible with the empirical evidence, perhaps supported by it, and more generally, that the theory of adjunction conforms to SMT in crucial aspects — maybe entirely. If so, another large category of phenomena shift from "unexplained" to "explained in a principled manner" — from (2i) to (2ii, iii).

Let's consider finally the place of Spell-Out S-O — that is, the choice of phases. We know that S-0 cannot apply at each stage of cyclic Merge. Relevant information may not yet be available. Suppose, for example, that Merge has constructed (see, OBJ), where the object OBJ = *that* or *what.* At this stage we do not know whether OBJ or *see* is spelled out in situ, or whether they move on overtly to be spelled out in a higher position. If they move on (either sometimes, or always), then S-0 plainly cannot apply at this stage. In a cyclic theory, we do not want to wait too long to determine whether they are spelled out in situ. Ideally, it should be at the next Merge. Suppose further that *see* is a root; a reasonable assumption, as noted earlier. Then the next Merge should also tell us what kind of element it is: the verb *see* or the noun *sight.* In the best case, then, the next Merge should yield (20), where a is the verbalizer *v* or the nominalizer *n:*

(20) {a, {see, OBJ}}

If a = *n,* then OBJ receives inherent Case and can be spelled out at this stage. Suppose it is *v.* We now have the conventional *v-VP* analysis. V raises to *v,* but we do not yet now whether it is spelled out.[69] *v* can be of various types. Suppose *v* is transitive. Then Agree holds

[67] See von Fintel and Iatridou (2001) for review, and some new cases.
[68] It also omits entirely the parasitic gap and other constructions that are the primary focus of Nissenbaum's work, because of their intricacy; but I think that an analysis similar to that proposed below may be possible here as well.
[69] 1 will omit further comment on head-raising, including the possibility, which I think is real, that it is part of O; see *DbP.* Also ignored are other possibilities, e.g., automatic raising of a to some position above V but below *v,* as in recent work of Masatoshi Koizumi, Howard Lasnik, and others.

between *v* and OBJ, and structural Case of OBJ can be assigned the value ACC, with <p-features of *v* valued by OBJ. Whether OBJ raises further at this stage depends on whether *v* has an OCC (=EPP) feature. Suppose it does. Then OBJ raises to SPEC-u, either above or below the externally-Merged subject SU that is required for convergence; see note 36. If below, then there will be an ultimate theta-theory violation, detectable at once. Proper positioning might be automatic under various assumptions: e.g., if the simultaneous satisfaction of properties of *v* involves an internal cyclic order, with raising of OBJ first, then "tucking in" of externally Merged SU. Recall that it cannot be required that every subsequent extension converges; only that there are convergent extensions.

S-O should apply as early as possible, for reasons already discussed. It cannot be before the *v* level, but it can be at this level, as long as only the sister of *v* is spelled out (in accord with PIC). Therefore *vP* is a phase. We understand PIC as before: the sister of the head is spelled out obligatorily; the fate of the edge — the head and its SPECs —is not determined until later; see note 19.

Suppose T is. merged with (20). If T is defective, SU may raise to SPEC-T, but it remains active. If T is complete, the next Merge is C. Though SU may raise to SPEC-T before C is merged, valuing and inactivating Case, this is really a reflex of the C-T relation, as we have seen. For the other two elements of the edge of *v, v* itself and the extra SPECs, we know their local destination by the C level, but in general not before: A'-movement is to SPEC-C, and the same appears to be true for V-topicalization (Holmberg 1999). Relevant considerations therefore converge to the conclusion that the next higher phase is CP (or *vP,* if T is defective).

Spell-Out applies at the phase level (by definition), and as discussed, all operations within the phase are in effect simultaneous. Furthermore, their applicability is evaluated at the phase level, yielding apparent countercyclic effects within the phase; see *DbP.* The phenomenon is illustrated most simply by A'-movement to SPEC-C, as in (21) (abstracting from effects of T-to-C, *dosuppoit,* and seem-to-T raising; *t* the trace (copy) of the *zvh-phiase, t'oihe):*

(21) (i) what C [he T [t [f seet]]]

(ii) to-whom C [he T [t seem [f to be intelligent]]

Applying cyclically, T first raises *he* to SPEC-T, skipping *wh-* in apparent violation of the Minimal link Condition MLC. After Merge of C, *wh-* is raised to SPEC-C, voiding the violation of MLC at the phase level.[70] Again, it is "as if' all operations are applying simultaneously at the phase level, as we should now expect.

Phase-level evaluation has other consequences. One is that if XP is raised to SPEC-a, it cannot remain there, or the external argument will not be accessible to the higher complete category with which it must agree (C-T or *v),* and the derivation will crash. We expect, then, that under successive-cyclic movement, w/i-phrases cannot be stranded in SPEC-P, though it

[70] Note that this provides still another reason to conclude that TP is not a phase.

could be strandable in SPEC-C. Both conclusions are apparently correct to my knowledge, no clear case of stranding in SPEC-P is known, but stranding in SPEC-C is a plausible analysis of the German uws-constructions. It also follows that if Object-Shift is to *SPEC-v,* it must then move on to a higher position; see *DbP* for discussion.[71]

We have good reason, then, to regard *vP* and CP (but not TP) as phases.[72]

Why should these be the phases, and the only ones? Ideally, phases should have a natural characterization in terms of IC: they should be semantically and phonologically coherent and independent. At SEM, *vP* and CP (but not TP) are prepositional constructions: *vP* has full argument structure, CP is the minimal construction that includes Tense and event structure,[73] and (at the matrix at least) force. At PHON, these categories are relatively isolable (in clefts, VP-movement, etc.). These properties do not, however, yield exactly the right distinctions: *vP* with *v* nontransitive is relatively isolated and is a domain for QR, though these cannot be phases for Spell-Out. Call these *weak phases.* Then the strong phases are those that have an EPP position as an escape hatch for movement, and are therefore the smallest constructions that qualify for Spell-Out.[74]

So far, we have remained reasonably close to the strong minimalist thesis (3); that is, within the categories (2ii, iii). That seems rather surprising. One reason is that even much weaker versions of the thesis are surprising (and difficult to discover and poorly understood for organic systems except at the simplest levels). Another is that nothing remotely similar would have been anticipated only a few years ago. How far this kind of analysis can proceed, or along what paths, one can only conjecture. Insofar as it can, the conclusions are of some significance, not only for the study of language itself.

[7] German, see McDaniel 1989. Suppose that SPEC-D is the position of overt *wh-* (and possibly focus). We might ask whether in such cases *v* has raised to T, so that the SPEC position does not intervene between T and EA. Many questions remain concerning Heavy-NP-Shift, Object-Shift, and overt and covert utfi-movement, all assumed here to move to SPEC-tf, though they have different properties.

[72] See note 16.

[73] Recall that T has these properties only as a reflex of C-T.

[74] On nontransitives, see Legate 1998. On the escape hatch in strong phases and parasitic gaps, see Nissenbaum 2000.

References

Agiiero, C. (2001). The Scope of Wh-phrases. Ms., MIT. Barss, A. (1986). *Chains and Anaphoric Dependence.* PhD dissertation, MIT. Boeckx, C. (2000). Quirky Agreement. *Studia Linguistica* 54.3, December. Boeckx, C. (2001). Experiencer-Intervention Across Languages. Ms., University of Connecticut. To appear in P. Pica and J. Rooryck, eids. *Syntax Yearbook* (Amsterdam: John Benjamins). Bo§kovic, Z. (1998a). W7i-phrases and Wi-movement in Slavic. Comparative Slavic Morphosyntax Workshop, Bloomington. Boskovid, Z. (1998b). Multiple a*-fronting and Economy of Derivation. In E. Curtis, J. Lyle, and G. Webster, eds., *Proceedings of the 16th West Coast Conference on Formal Linguistics,* 49-63. Stanford: CSLI Publications. [Distributed by Cambridge University Press.]

Brody, M. (1995). *Lexico-Logical Form: a Radically Minimalist Theory.* Cambridge MA: MIT Press. Carstairs-McCarthy (1999). *The Origins of Complex Language.* Oxford: Oxford University Press.

Cherniak, C. (1995). Neural Component Placement. *Trends in Neuroscience* 18,522-27.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge MA: MIT Press. Chomsky, N. (1970). Remarks on Nominalization. In R. Jacobs and P. Rosenbaum, eds. *Readings in English Transformational Grammar.* Waltham MA, Ginn and Co. Reprinted in Chomsky, *Studies on Semantics in Generative Grammar* (1972). The Hague: Mouton.

Chomsky, N. (2000a). *New Horizons in the Study of Language and Mind.* Cambridge UK: Cambridge University Press. Chomsky, N. (2000b). Minimalist Inquiries *(MI).* In Martin et al., eds. *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik.* Cambridge MA: MIT Press. Chomsky, N. (2001). Derivation by Phase *(DbP).* In Kenstowicz, ed. *Ken Hale: A life in Language.* Cambridge MA: MIT Press. Chomsky, N., R. Huybregts and H. van Riemsdijk (1982). *The Generative Enterprise.* Dordrecht: Foris.

Chung, S. (1998). *The Design of Agreement.* Chicago: University of Chicago Press.

Collins, C. (1993). *Topics in Ewe Syntax.* PhD dissertation, MIT. Distributed by MIT Working Papers in Linguistics. Collins, C. (2001). Eliminating Labels. In *MIT Occasional Papers in Linguistics 20* (this volume). To appear in Epstein, Samuel David and T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program* (Blackwell, forthcoming 2002). Epstein, S. (1999). Un-Principled Syntax: the Derivation of Syntactic Relations. In Epstein and Homstein, eds. *Working Minimalism.* Cambridge MA: MIT Press. Epstein, S. and N. Homstein, eds. (1999). *Working Minimalism.* Cambridge MA: MIT press. Fintel, K. von, and S. Iatridou (2001). On the Interaction of Modals, Quantifiers, and *If*-Clauses. Ms., MIT.

Fox, D. (2000). *Economy and Semantic Interpretation.* Cambridge MA: MIT Press. Fox, D. (2001). Antecedent Contained Deletion and the Copy Theory of Movement. Ms., Harvard. Fox, D., and J. Nissenbaum (1999). Extraposition and the Nature of Covert Movement. Ms., Harvard.

Frampton, J., S. Gutmann, J. Legate, and C. Yang (1999). Remarks on DBP, Agreement, and Intervention. Ms., Northeastern University and MIT.

Frampton, J., and S. Gutmann (2000). Crash-Proof Syntax. Ms., Northeastern University.

Freidin, R., and J-R Vergnaud (2001). Exquisite Connections: Some remarks on the evolution of linguistic theory. Ms., Princeton, USC. Fukui, N. (2001). Phrase Structure. In M. Baltin and C. Collins, eds., *The Handbook of Contemporary Syntactic Theory.* Maiden MA, Oxford: Blackwell. Gil, D. (1987). On the Scope of Grammatical Theory. In S. Mogdil and C. Mogdil, eds. *Noam Chomsky: Consensus and Controversy* (Philadelphia: Taylor & Francis). Guasti, M. T., and L. Rizzi (1999). Agreement and Tense as Distinct Syntactic Positions: Evidence from Acquisition. Ms., Siena. Hiraiwa, K. (2000). Multiple Agree and the Defective Intervention Constraint in Japanese. Ms., MIT. Holmberg, A. (1999). Remarks on Holmberg's Generalization. *Studia Linguistica* 53.1,1-39, April.

Homstein, N. (1999). Movement and Control. *Linguistic Inquiry* 30.69-96. Iatridou, S. (1993). On Nominative Case Assignment and a Few Related Things. C. Phillips, ed., *MIT Working Papers in Linguistics 19: Papers on.Case & Agreement II,* 175-96.

Iatridou, S. (2001). Temporal Existentials? Ms., MIT.

Jenkins, L. (2000). *Biolinguistics.* Cambridge UK: Cambridge University Press. Kayne, R. (1981). Unambiguous paths. In R. May and J. Koster, eds., *Levels of Syntactic Representation.* Dordrecht: Foris.

Kayne, R. (1994). *The Asymmetry of Syntax.* Cambridge MA: MIT Press. Kenstowicz, M., ed. (2001). *Ken Hale: A Life in Language.* Cambridge MA: MIT Press. Lasnik, H. (1999a). *Minimalist Analysis.* Maiden MA, Oxford: Blackwell. Lasnik, H. (1999b). On the Locality of Movement: Formalist syntax position paper. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, and K. Wheatley, eds. *Functionalism and Formalism in linguistics,* vol. I. Amsterdam: John Benjamins. Lasnik, H. (1999c). Chains of Arguments. In S. Epstein and N. Homstein, eds. *Working Minimalism.* Cambridge MA: MIT Press. Lebeaux, D. (1988). *Language Acquisition and the Form of Grammar.* PhD Dissertation, U. Mass Amherst. Revised and extended version (2000). Amsterdam: John Benjamins Lebeaux, D. (1999). Where does the Binding Theory Apply? Ms., NEC Research Institute. Legate, J. (1998). Verb Phrase Types and the Notion of a Phase. Ms., MIT. Leiber, J. (2001). Turing and the fragility and insubstantiality of evolutionary explanations: a puzzle about the unity of Alan Turing's work with some larger implications. *Philosophical Psychology* 14.1. Martin, R, and J. Uriagereka (2000). Introduction: Some Possible Foundations for the Minimalist Program. In Martin et al., eds., *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik.* Cambridge MA: MIT Press. Martin, R, D. Michaels and J. Uriagereka, eds. *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik.* Cambridge MA: MIT Press. Maynard Smith, J. (1998). *Shaping Life: Genes, Embryos and Evolution.* London: Weidenfeld & Nicolson. Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, L. Wolpert (1985). Developmental Constraints in Evolution. *Quarterly Review ofBiology* 60.3,265-87.

McDaniel, D. (1989). Partial and Multiple VWi-movement. *Natural Language and Linguistic Theory* 7,565-604.
McCloskey, J. (2000a). Quantifier float and Wh-movement in an Irish English. *Linguistic Inquiry* 31.1:57-84.
McCloskey, J. (2000b). The Distribution of Subject Properties in Irish. Ms., UC Santa Cruz. To appear in S. Dubinsky and W. Davies, eds., 77K *Role of Grammatical Functions in Transformational Syntax.* Dordrecht Kluwer. McCloskey, J. (2000c). Resumption, Successive Cyclidty, and the Locality of Operations. Ms., UC Santa Cruz.
Miller, G., and N. Chomsky (1963). Finitary Models of Language Users. In R. Luce, R. Bush, and E. Galanter, eds. *Handbook of Mathematical Psychology* II, chap. 13,419-492. New York, Wiley.
Nissenbaum,J. (2000). *Investigations of Covert Phrase Movement.* PhD Dissertation, MIT.
Pesetsky, D. (2000). *Phrasal Movement and its Kin.* Cambridge MA: MIT Press. Pesetsky, D., and E. Torrego (2001). In Kenstowicz, ed., *Ken Hale: A Life in Language.* Cambridge MA: MIT Press.
Petirto, L, R. Zatorre, K. Gauna, E. Nikelski, D. Dostie, A. Evans (2000). Speech-like cerebral activity in profoundly deaf people while processing signed languages: Implications for the neural basis of human language. *Proceedings of the National Academy of Sciences USA* 97.25:13961-13966, December 5. Ramus, F., M. Hauser, C. Miller, D. Morris, J. Mehler (2000). Language Discrimination by Human Newboms and by Cotton-Top Tamarin Monkeys. *Science.* Richards, N. (1997). *What Moves Where in Which Language.* PhD Dissertation, MIT. Distributed by MIT Working Papers in Linguistics.
Roeper, T. (2000). Universal Bilingualism. *Bilingualism: Language and Cognition 2.*
Saddy, J. (1991). *Investigations into Grammatical Knowledge.* PhD Dissertation, MIT.
Stewart, I. (1998). *Life's Other Secret.* New York: Wiley. Uriagereka,J. (1998). *Rhyme and Reason.* Cambridge MA: MIT Press. Uriagereka, J. (1999). "Multiple Spell-Out." In Epstein and Horstein, eds. *Working Minimalism.* Cambridge MA: MIT Press.
Yang, C. (2000). *Knowledge and Learning in Natural Language.* PhD Dissertation, MIT. Distributed by MIT Working Papers in Linguistics.

BEYOND EXPLANATORY ADEQUACY

by

Noam Chomsky

&

ELIMINATING LABELS

by

Chris Collins

MIT Occasional Papers in Linguistics
Number 20

MIWE

2001

Distributed by
MIT Working Papers in Linguistics
E39-245, MIT, Dept. of Linguistics
77 Massachusetts Avenue
Cambridge, MA 02139 USA
mitwpl@mit.edu //
http://mit.edu/mitwpl/