On Phases*
NOAM CHOMSKY (MIT)
26 September 2005

I would like to review some recent and ongoing work in the general framework of the so-called "Minimalist Program" (MP), which addresses an array of concerns that are a core part of the traditional study of language, assuming a different form within the biolinguistic perspective that began to take shape fifty years ago. I will presuppose familiarity with recent publications, while recalling some conclusions that seem pertinent to proceeding along lines they suggest. One major stimulus to the development of the MP was a personal letter by Jean-Roger Vergnaud in 1977, unpublished but famous in the field, which initiated very extensive and productive inquiries that remain central to the study of language to the present, including the considerations here.[1]

The traditional concerns have to do with the properties that are specific to human language, that is, to the "faculty of language" FL. To borrow Jespersen's formulation eighty years ago, the goal is to unearth "the great principles underlying the grammars of all languages" with the goal of "gaining a deeper insight into the innermost nature of human language and of human thought." The biolinguistic perspective views FL as an "organ of the body," one of many subcomponents of an organism that interact in its normal life. From this perspective, the closest approximation to the informal notion "language" is a state of FL, an I-language. UG is the theory of the initial state of FL, virtually shared; in terms of traditional concerns, the theory of the distinguishing features of human language.

For any such system, we can identify three factors that enter into its growth and development: (I) external data; (II) genetic endowment (for language, the topic of UG); (III) principles of structural architecture and developmental constraints that are not specific to the organ under investigation, and may be organism-independent. Factor (II) interprets data as linguistic experience, not a trivial matter; it constructs part of what ethologists called the organism's *Umwelt*.[2] And it sets the general course of growth and development within a narrow range. The particular path is determined by experience (among other factors that we can ignore here). The questions arise in principle for any organic system, and have been raised since the early days of modern biology. The only general questions concerning them have to do with feasibility, not legitimacy.

These questions arose in the early days of the biolinguistic perspective, which converged quickly with efforts to develop generative grammar, rooted in different concerns. The matter was brought up in print in the course of critical discussion of prevailing assumptions in the "behavioral sciences": basically, the belief that language, like all behavior, is shaped by association, conditioning, and simple induction, along with "analogy," and perhaps also "general learning procedures" of some unspecified sort. While such notions were understandable "in the context of eighteenth century struggles for scientific naturalism,...there is surely no reason today for taking seriously a position that attributes a complex human achievement entirely to months (or at most years) of experience, rather than to millions of years of evolution or to principles of neural organization that may be more deeply grounded in physical law," a stand that would furthermore

[1] Letter, April 17, 1977; Vergnaud (1985). I would like to thank Robert Freidin, Samuel Epstein, and Eric Reuland, among others, along with participants in a Fall 2004 seminar at MIT, for valuable comments on an earlier draft.
[2] Obviously, this is an extreme abstraction. Each stage of development, along various paths, contributes to interpretation of data as experience.

take humans to be "unique among animals."[3] The three factors, essentially. The issues were not much discussed in public, partly because the third factor was at the remote horizons of inquiry (as commonly in other domains), but also because of the much more urgent concern of justifying even the legitimacy of the generative grammar/biolinguistic amalgam as a way of studying language at a time when even the second factor was considered highly contentious or worse, and it was commonly assumed that languages can "differ from each other without limit and in unpredictable ways,"[4] or could vary at most within limited constraints on patterns and structures.

UG is said to meet the condition of "explanatory adequacy" insofar as it is a true theory of the genetic endowment, and thus, under appropriate idealizations, maps experience to attained I-languages -- abstracting here from effects of the third factor. We can regard an explanation of some property of language as *principled*, to the extent that current understanding now reaches, insofar as it can be reduced to the third factor and to conditions that language must meet to be usable at all – specifically, conditions coded in UG that are imposed by organism-internal systems with which FL interacts. Insofar as properties of I-languages can be given a principled explanation, in this sense, we move to a deeper level of explanation, beyond explanatory adequacy.

Though these terms are not used in other domains – say, bee communication, or the mammalian visual system – investigation adopts effectively the same categories. In the study of language as in other domains, it is uncontroversial that search for explanatory adequacy not only does not await achievement of descriptive adequacy, but rather contributes to that goal and even to discovery of the nature of the task, by clarifying the true nature of the object of inquiry (I-language) and of descriptive adequacy. It is no less a truism that the same relation holds between both inquiries and the search for principled explanation.

Concern for principled explanation is often framed in methodological terms as the attempt to keep to simple taxonomies or generative systems without excessive redundancy and other unattractive properties. Sometimes methodological considerations, largely intuitive, can be reframed within the biolinguistic perspective as empirical hypotheses that can be investigated in other domains as well. For computational systems such as language, we naturally hope to discover concepts of computational efficiency that carry us beyond explanatory adequacy, and to investigate how these relate to principles of a more general character that may hold in other domains and for other organisms, and may have deeper explanations.

For reasons that need not be reviewed,[5] the crystallization of the Principles-and-Parameters (P&P) program removed some serious conceptual barriers to exploring the possibility of principled explanation. Since that time, it has been the subject of extensive inquiry, from various approaches, with sufficient progress and convergence, many researchers have felt, to identify a reasonably integrated sub-discipline that focuses on these issues, the minimalist program MP.[6]

---

[3] Chomsky (1965), 59.

[4] Joos's (1957) summary of the guiding "Boasian" tradition of American linguistics.

[5] For some discussion, see Chomsky (2005).

[6] See Brody (1995), Chomsky (1993, 1995), and much subsequent work. For illuminating comment on the general project, see Freidin and Vergnaud (2001). Despite repeated clarification, MP is often taken to be a hypothesis about language or a new approach to language, displacing earlier ones. It is neither. Furthermore, as again repeatedly stressed, the program is theory neutral: whatever one's conception of UG, one can be interested in principled explanation (MP), or not. And if so, essentially the same questions will arise.

Adopting the P&P framework, I will assume that one element of parameter-setting is assembly of features into lexical items (LIs), which we can take to be atoms for further computation and the locus of parameters, sweeping many complicated and important questions under the rug.

It has been a useful guide for research to consider an extremely far-reaching thesis – the "strong minimalist thesis" SMT – which holds that language is an optimal solution to interface conditions that FL must satisfy; that is, language is an optimal way to link sound and meaning, where these notions are given a technical sense in terms of the interface systems that enter into the use and interpretation of expressions generated by an I-language. If SMT held fully, which no one expects, UG would be restricted to properties imposed by interface conditions. A primary task of the MP is to clarify the notions that enter into SMT and to determine how closely the ideal can be approached. Any departure from SMT – any postulation of descriptive technology that cannot be given a principled explanation – merits close examination, to see if it is really justified. The more fully principled explanation can be achieved, the better we understand the nature of FL; and the better we can formulate, perhaps even constructively investigate, the problem of how FL evolved, apparently appearing quite recently in a small breeding group of which all contemporary humans are descendants, part of a "great leap forward" in human intellectual and moral faculties, as some paleoanthropologists term this development. Evidently, inquiry into evolutionary origins becomes more feasible the less special structure is attributed to UG: that is, the more we can proceed beyond explanatory adequacy.[7]

It is hardly necessary to add that the conditions that enter into principled explanation, in this sense, are only partially understood: we have to learn about the conditions that set the problem in the course of trying to solve it. The research task is interactive: to clarify the nature of the interfaces and optimal computational principles through investigation of how language satisfies the conditions they impose – optimally, insofar as SMT holds. This familiar feature of empirical inquiry has long been taken for granted in the study of the sensory-motor interface (SM). Inquiry into acoustic and articulatory phonetics takes cues from what has been learned about phonological features and other such properties in I-language research and seeks SM correlates, and any discoveries then feed back to refine I-language inquiry. The same should hold, no less uncontroversially, at the semantic/conceptual-intentional interface (C-I). And it should also hold for third factor properties. We do not know a priori, in more than general terms, what are the right ways to optimize, say, neural networks; empirical inquiry into such matters is interactive in the same ways.

A further question is whether the contribution of the two interfaces to principled explanation is symmetrical. It is well-known that language is in many ways "poorly designed" for communicative efficiency: apart from such ubiquitous phenomena as ambiguity, garden paths, etc., the core property of language – recursive embedding with nested dependencies – leads to exponential memory growth and therefore has to be avoided in language use, giving it something of the character of paratactic constructions. Languages have various devices to overcome the problems.[8] These devices might be close to or even beyond the SM interface. Some of them are used to overcome prosodic difficulties, as in familiar examples of the "house that Jack built" variety. Others yield "rearrangements" near the SM interface that violate crossing constraints and have other properties that indicate that they are not operations of the narrow syntax. PP-extraposition is a likely case, to which we return.[9]

[7] For some discussion and sources, see Chomsky (2005).
[8] See, e.g., O'Neil (1977), and sources cited. Also discussion in Miller and Chomsky (1963).
[9] See Chomsky (1995), chap. 4, 7.3. Such considerations suggest that the freezing principle of Wexler and Culicover (1980) should be generalized to immobilization of the full extraposed phrase, within narrow syntax. On apparent counterexamples with

It might be, then, that there is a basic asymmetry in the contribution to "language design" of the two interface systems: the primary contribution to the structure of FL may be optimization of mapping to the C-I interface. Such ideas have the flavor of traditional conceptions of language as in essence a mode of expression of thought, a notion restated by leading biologists as the proposal that language evolved primarily as a means of "development of abstract or productive thinking" and "in symbolizing, in evoking cognitive images" and "mental creation of possible worlds" for thought and planning, with communicative needs a secondary factor in language evolution.[10] If these speculations are on the right track, we would expect to find that conditions imposed by the C-I interface enter into principled explanation in a crucial way, while mapping to the SM interface is an ancillary process. If so, we might discover that SMT is satisfied by phonological systems that violate otherwise valid principles of computational efficiency, while doing the best it can to satisfy the problem it faces: to map to the SM interface syntactic objects generated by computations that are "well-designed" to satisfy C-I conditions. There is, I think, empirical evidence that something like that might be correct. But, again, the questions can only be answered by interactive research in many dimensions. Such questions are worth keeping in mind, even though they are at the periphery of current empirical study.

Suppose we assume SMT to be true, and see how far we can go towards accommodating properties of language, identifying places in the argument where assumptions are introduced that require independent research, and where the quest fails, for now at least.

In the early '90s, Howard Lasnik and I sketched our best understanding of what UG might be, adopting the familiar EST/Y-model, within the P&P framework (Chomsky and Lasnik 1993). A great deal of work since then, including ours, has been devoted to investigating stipulated components of that model to determine whether the phenomena for which they were designed could be derived by keeping more closely to SMT – or perhaps even better, with richer empirical coverage.

Two linguistic levels are assumed to be indispensable (though it is more than truism): the interface levels that are accessible to SM and C-I. These language-external though organism-internal systems have their own properties, which for present purposes we can assume to be independent of language, suppressing important questions that merit independent inquiry. The EST/Y-model postulated three additional internal levels, each with specific properties: d-structure, s-structure, and LF.[11] Furthermore, each of the five postulated levels is generated by cyclic/compositional operations, which are highly redundant, covering much the same ground. That sharp departure from SMT raised the question whether all of this technology can be reduced to a single cycle, dispensing with all internal levels. An even better result would be that the three internal levels are not only dispensable, but literally unformulable. That seems possible, if we examine more closely the generative procedure that constructs interface representations. I will briefly review (and somewhat restate) the line of argument in earlier papers.

---

semantic consequences (involving "late Merge" and Antecedent-contained-deletion), see Chomsky (2004), suggesting that these are actually generated in-situ. For additional evidence in support of this conclusion, see Szczegielniak (2004), showing that raising relatives have the same crucial properties, barring the possibility of QR-style covert-raising of the nominal head of the relative clause as proposed in the papers discussed in Chomsky (2004).

[10] Nobel laureates Salvador Luria and François Jacob. See Chomsky (2005) for context and sources.

[11] Some confusion has been caused by recent use of the term "LF" to refer to the C-I interface itself, thus departing from its definition within EST as the output of (overt and covert) syntactic operations and the input to rules of semantic interpretation that map LF to C-I. I am keeping here to the original sense, within the EST/Y-model. There is no issue beyond terminology.

As has long been recognized, the most elementary property of language – and an unusual one in the biological world – is that it is a system of discrete infinity consisting of hierarchically organized objects.[12] Any such system is based on an operation that takes *n* syntactic objects (SOs) already formed, and constructs from them a new SO. Call the operation *Merge*. Unbounded Merge or some equivalent is unavoidable in a system of hierarchic discrete infinity, so we can assume that it "comes free," in the present context.

Of course, the operation does not "come free" in human evolution. Rather, its emergence was a crucial event, so far without any explanation. Suppose that some ancestor, perhaps about 60,000 years ago, underwent a slight mutation rewiring the brain, yielding unbounded Merge. Then he or she would at once have had available an infinite array of structured expressions for use in thought (planning, interpretation, etc.), gaining selectional advantages transmitted to offspring, capacities that came to dominate, yielding the dramatic and rather sudden changes found in the archeological record. Speculation, of course, as are all such stories, but about the simplest one imaginable, consistent with what little is known, and presupposed in some form (often tacitly) in all speculations about the matter. Since the integration of language precursors into FL together with innovation of unbounded Merge would have been sudden (in evolutionary time), effects of path-dependent evolution and other complexities that underlie the logic of Jacobian "bricolage" might be secondary phenomena, and evolution to a form of FL approaching SMT, not too surprising.

Though study of the evolution of language is in its infancy,[13] consideration of the nature of the problem can be of some help in thinking about the core problems of study of language. It is hardly controversial that FL is a common human possession apart from pathology, to an approximation so close that we can ignore variation. Evidently, study of evolution of language is not concerned with developments since the "great leap forward" and the trek from Africa: say, with the large-scale effects of the Norman Conquest on English. It should be equally uncontroversial that investigation of explanatory adequacy and beyond will try to abstract from such developments to the extent that understanding permits, thus putting to the side many topics of great interest to anyone concerned with language.[14]

Returning to SMT, arguably restriction of computational resources limits *n* for Merge to two, as Luigi Rizzi suggests, thus yielding the "unambiguous paths" structure postulated by Kayne (1981). There are other conditions that conspire to this conclusion: minimal search within a probe-goal framework, for example. Perhaps also interface conditions: at the SM interface, requirements of linearization, perhaps along the lines of Kayne's LCA; at the C-I interface, conditions of predicate-argument structure and others. Let us assume the limitation to be accurate, at least sufficiently so that we can take it as a good first approximation.

A natural requirement for efficient computation is a "no-tampering condition" NTC: Merge of X and Y leaves the two SOs unchanged. If so, then Merge of X and Y can be taken to yield the set {X, Y}, the

[12] Hierarchy is automatic for recursive operations, conventionally suppressed for those that merely enumerate a sequence of objects.

[13] And might not proceed much beyond, some leading evolutionary biologists believe. See particularly Lewontin (1990). His chapter in an updated 1995 edition is no less skeptical about the prospects.

[14] Topics that may, however, have bearing on growth of language in the individual, and indirectly even on language evolution. As an example, Yang (2002), adapting ideas of Morris Halle's, has shown that contrary to widely-held beliefs, even in such marginal corners of English as irregular verbs, children appear to impose a principled system of generation, assigning the elements to rule-governed categories, thus providing evidence about the role of extra-linguistic principles of efficiency in determining the nature of I-languages attained.

simplest possibility worth considering. Merge cannot break up X or Y, or add new features to them. Therefore Merge is invariably "to the edge" and we also try to establish the "inclusiveness principle," dispensing with bar-levels, traces, indices and similar descriptive technology introduced in the course of derivation of an expression. It seems that this desideratum of efficient computation can also be met within narrow syntax at least, with apparent departures that have a principled explanation; and that it sometimes even yields superior empirical results, in one well-studied case, with regard to reconstruction effects. There is an ample literature on these matters, and I will assume it to be more less on the right track.

Note that SMT might be satisfied even where NTC is violated – *if* the violation has a principled explanation in terms of interface conditions (or perhaps some other factor, not considered here). The logic is the same as in the case of the phonological component, already mentioned.

A more complex alternative, consistent with NTC, is that Merge forms the pair <X, Y>. The underlying issue is whether linear order plays a role in narrow syntax and mapping to C-I, or whether it is restricted to the phonological component, motivated by interface conditions at SM. The latter assumption has guided a good deal of research since Reinhart (1979), and while the issue is far from settled, there seems to me good reason to suppose that the simpler assumption can be sustained: that order does not enter into the generation of the C-I interface, and that syntactic determinants of order fall within the phonological component.

Suppose that a language has the simplest possible lexicon: just one LI, call it "one." Application of Merge to the LI yields{one}, call it "two" Application of Merge to{one}yields{{one}}, call it "three." Etc. In effect, Merge applied in this manner yields the successor function. It is straightforward to define addition in terms of Merge(X, Y), and in familiar ways, the rest of arithmetic. The emergence of the arithmetical capacity has been puzzling ever since Alfred Russell Wallace, the co-founder of modern evolutionary theory, observed that the "gigantic development of the mathematical capacity is wholly unexplained by the theory of natural selection, and must be due to some altogether distinct cause," if only because it remained unused. It may, then, have been a side product of some other evolved capacity (not Wallace's conclusion), and it has often been speculated that it may be abstracted from FL by reducing the latter to its bare minimum. Reduction to a single-membered lexicon is a simple way to yield this consequence.[15]

For an LI to be able to enter into a computation, merging with some SO (and automatically satisfying SMT), it must have some property permitting this operation. A property of an LI is called a *feature*, so an LI has a feature that permits it to be merged. Call this the *edge-feature* (EF) of the LI. If an LI lacks EF, it can only be a full expression in itself; an interjection. When merged with a syntactic object SO, LI forms {LI, SO}; SO is its *complement*. The fact that Merge iterates without limit is a property at least of LIs – and optimally, only of LIs, as I will assume. EF articulates the fact that Merge is unbounded, that language is a recursive infinite system of a particular kind. What kind? That depends on further specification of EF, which, like any other feature of UG – say, sonorant – has to be defined. Its definition, to which we return, is an empirical hypothesis about the subcategory of recursive infinite systems that language constitutes.

Reliance on iterable Merge as the sole computational operation of narrow syntax eliminates, as unformulable, the notions d- and s-structure, and three of the compositional operations of EST: those that form d-structure and map it to s-structure, and then on to LF. It also revives, in far more elementary terms, the notion of generalized transformation of the earliest work in generative grammar in the 1950s. A great

[15] Dissociations in aphasia and performance, which have been known for many years, tell us little in this regard, contrary to much discussion, because they reflect many different factors. Similarly, dissociation of reading from other language performance has never been taken to entail that there is a "reading faculty" independent of the faculty of language.

deal of descriptive capacity is lost by this optimal assumption, including all processes and principles associated with d- and s-structure, bar-level distinctions in phrase structure (N vs. N', etc.), and much else. Since this descriptive technology has been widely and productively used, it is a significant empirical task to show that it is dispensable – to show, that is, that it was misconceived, and should be abandoned. I will assume that success in this endeavor has been sufficient so that it is reasonable to assume it for our purposes here.

Suppose that X and Y are merged (for expository purposes, think of Y as merged to X). Either Y is not part of X (*external Merge*, EM) or Y is part of X (*internal Merge*, IM).[16] In both cases, Merge yields {X, Y}. IM yields two *copies* of Y in {X, Y}, one external to X, one within X.[17] IM is the operation *Move* under the "copy theory of movement," which is the null hypothesis in this framework, required by strict adherence to the NTC. Unless there is some stipulation to the contrary, which would require sufficient empirical evidence, both kinds of Merge are available for FL and IM creates copies.

If the means of language are fully exploited by the interface systems, in accord with a reasonable interpretation of SMT, then we would expect the two types of Merge to have different effects at the interfaces. At the phonetic interface, they obviously do; IM yields the ubiquitous displacement phenomenon. At the semantic interface, the two types of Merge correlate well with the duality of semantics that has been studied within generative grammar for almost 40 years, at first in terms of "deep and surface structure interpretation" (and of course with much earlier roots). To a large extent, EM yields generalized argument structure (theta roles, the "cartographic" hierarchies,[18] and similar properties); and IM yields discourse-related properties such as old information and specificity, along with scopal effects. The correlation is reasonably close, and perhaps would be found to be perfect if we understood enough – an important research topic.

The two available types of Merge have been treated very differently since the early days of modern generative grammar. But that is a historical residue; we have to ask whether it is accurate.

It has always been presupposed without comment that EM comes free: no one has postulated an "EPP property" for EM or stipulated that it satisfies the NTC. IM, in contrast, has been regarded (by me, in particular) as a problematic operation, an "imperfection" of language that has to be postulated as an unexplained property of UG unless it can be motivated in some principled way. The displacement property of natural language is simply a fact: expressions are commonly pronounced in one place and interpreted in others as well. There have been many ideas as to how to capture that fact, transformational operations being one. My own view had always been that the alternatives might turn out to be close to notational variants, and that any decision among them would have to be teased out by subtle empirical evidence. A few years ago, it became clear that this is a misunderstanding. IM (= Move, with the "copy theory") is as free as EM; it can only be blocked by stipulation. The *absence* of the operation would be an imperfection that has to be explained, not its use in deriving expressions. It follows that any alternative device to deal with the displacement property and the duality of semantics requires double stipulation: to ban IM, and to

[16] "Part of" here means "term of," in the technical sense: a term of the SO α is a member of α or of a term of α.

[17] There has been much misunderstanding since the copy theory was proposed in Chomsky (1993), modifying earlier conceptions of movement by eliminating trace and indexing in favor of the NTC: that is, leaving the moved element unaffected instead of replacing it by an indexed trace (indexing is now superfluous, under identity). It has sometimes been supposed that a new "copy" is created, then inserted in the position of the moved element – all unnecessary -- and an alternative has been proposed in terms of "remerge," which is simply a notation for the copy theory as originally formulated in the most elementary terms.

[18] See Cinque (1999); Cinque, ed. (2002); Rizzi, ed. (2004); Belletti, ed. (2004).

justify the new device. The proposal faces a serious empirical burden of proof, unlike the core (and over time, much refined) principle of transformational grammar.

We thus expect language to use IM rather than other mechanisms that can be devised to express semantic properties apart from generalized argument structure; largely true, it appears, perhaps better than that. Note that this line of argument is an adaptation to the C-I interface of the reasoning familiar at the SM interface, already discussed. Here too, we may regard these observations as an empirical hypothesis about the nature of the C-I system, to be investigated in language-independent terms. The hypothesis is that C-I incorporates a dual semantics, with generalized argument structure as one component, the other being discourse-related and scopal properties. Language seeks to satisfy the duality in the optimal way, EM serving one function and IM the other, avoiding additional means to express these properties.

Each SO generated enters into further computations. Some information about the SO is relevant to these computations. In the best case, a single designated element should contain all the relevant information: the *label* (the item "projected" in X'-theories; the *locus* in the label-free system of Collins 2002). The label selects and is selected in EM, and is the *probe* that seeks a *goal* for operations internal to the SO: Agree or IM. A natural interpretation of the notion "edge" can capture some of the properties of "tucking in" in the sense of Richards (2001), taking the "edge" to be the position as close as possible to the probe – a literal violation of NTC, but arguably a principled one, hence consistent with SMT.

We therefore have two syntactic relations: (A) set-membership, based on Merge, and (B) probe-goal relations. Assuming composition of relations, (A) yields the notions *term-of* and *dominate*. These seem to be the minimal assumptions about the available relations. If we add "sister-of," then composition will yield c-command and identity (the latter presumably available independently). Whether c-command plays a role within the computation to the C-I interface is an open question. I know of no clear evidence that it does, so will keep to the relations that seem unavoidable, set-membership and probe-goal.

It has always been assumed that c-command relations function in Binding Theory BT, which would mean, if the above is correct, that BT is at the outer edge of the C-I interface (as suggested, in essence, in Chomsky and Lasnik 1993). It is not entirely clear, however, that c-command is actually required here, in addition to set-membership and probe-goal. Condition (C) could be formulated as a probe-goal relation, taking the c-commanding pronoun X to be the label of {X, SO}, hence a probe (along lines to which we turn below, though with complications and consequences I will put aside here). As for Condition (A), suppose we adopt the framework of Reuland (2001). Then the most important case is bare subject-oriented reflexives R that satisfy locality conditions, and are thus plausibly to be understood as within the I-language. In a structure of the form {SPEC, {H, ...R....}}, with R c-commanded and bound by SPEC, R could be taken to be the goal probed by H, and thus only indirectly bound by SPEC; hence a case of Agree, not c-command. The crucial empirical test is long-distance agreement: a structure of the form [H...XP...R], where H and XP agree, XP does not c-command R, and R is in the minimal search domain of the probe H; a sentence of the basic form (1):

(1) it became [[introduced a man] for R (self)]

Here "a man" does not c-command R, but both "a man" and R are goals of the probe that heads the construction. Reuland points out[19] that in an interesting range of such cases the reflexive must have the

[19] Pc. His examples are from Norwegian and Icelandic.

bare form R, meaning it is in an agreement (probe-goal) relation with H, though not c-commanded by its antecedent XP. If so, then the core case of Condition (A) does not involve c-command, but rather Agree. C-command may turn not to be an operative relation for Condition (A), which would support the view that the only relations are the inescapable ones: set-membership and probe-goal.

I will take over here the probe-goal framework of earlier work, including Ken Hiraiwa's theory of Multiple-Agree.[20] The probe agrees with goals in its domain as far as a goal with no unvalued features, which blocks further search (intervention). In the simplest case of two-membered probe-goal match (say [φ-features]-N), intrinsic features of the goal value those of the probe, and also value the structural Case feature of the goal (in a manner determined by the probe). Generalizing to Multiple-Agree (e.g., probe-[participle sequence]-N), features of the goal (including the structural Case, sometimes visible as in Icelandic) value all matched elements (probe, participles), and the option of raising the goal may or may not be exercised[21] -- and if exercised will be all the way to the probe, matters to which we return.

As noted, iterated Merge incorporates the effects of three of the EST/Y-model compositional cycles, while eliminating d- and s-structure. Still unaccounted for are the cyclic/compositional mappings to the phonetic and semantic interfaces. These too are incorporated, and the final internal level LF is eliminated, if at various stages of computation there are Transfer operations: one hands the SO already constructed to the phonological component, which maps it to the SM interface ("Spell-Out"); the other hands SO to the semantic component, which maps it to the C-I interface. Call these SOs *phases*. Thus SMT entails that computation of expressions must be restricted to a single cyclic/compositional process with phases. In the best case, the phases will be the same for both Transfer operations. To my knowledge, there is no compelling evidence to the contrary.[22] Let us assume, then, that the best-case conclusion can be sustained. It is also natural to expect that along with Transfer, all other operations will also apply at the phase level, as determined by the label/probe. That implies that IM should be driven only by phase heads. Empirical evidence to which we return supports this conclusion.

For minimal computation, as soon as the information is transferred it will be forgotten, not accessed in subsequent stages of derivation: the computation will not have to look back at earlier phases as it proceeds, and cyclicity is preserved in a very strong sense.[23] Working that out, we try to formulate a *phase-impenetrability condition* PIC, conforming as closely as possible to SMT. I will again assume that the literature about this is more or less on target (including Chomsky 2001, 2004; and Nissenbaum 2000), though with modifications.[24]

Note that for narrow syntax, probe into an earlier phase will almost always be blocked by intervention effects. One illustration is agreement into a lower phase without intervention in experiencer constructions in which the subject is raised (voiding the intervention effect) and agreement holds with the nominative

[20] Chomsky (2004). See Hiraiwa (forthcoming) for discussion of these and many related questions along lines rather similar to those adopted here. Also Boeckx (2004b).
[21] As noted in Chomsky (2004), note 51, generalization to Multiple-Agree overcomes problems in Chomsky (2001) discussed by Frampton et al. (2000), who develop an alternative approach in this and later papers; see Frampton and Guttman (2004) for recent extensions. The same processes hold elsewhere, e.g., in Hindi multiple-gender agreement; see Boeckx (2004a), Hiraiwa (forthcoming).
[22] The optimal conclusion is, in fact, required, under the theory of adjuncts outlined in Chomsky (2004).
[23] On various notions of cyclicity, often confused, see Lasnik (in press). Note that the considerations discussed here suggest a principled explanation for the successive-cyclic property of A'-movement.
[24] To mention one, in the analysis developed below we can dispense with the "next higher phase" condition of my earlier papers cited.

object of the lower phase (Icelandic).[25] It may be, then, that PIC holds only for the mappings to the interface, with the effects for narrow syntax automatic.

The next question is: What are the phases? I will pursue the suggestion in Chomsky (2004) that they are CP and v*P, where C is shorthand for the region that Rizzi (1997) calls the "left periphery," possibly involving feature spread from fewer functional heads (maybe only one); and v* is the functional head associated with full argument structure, transitive and experiencer constructions, and is one of several choices for v, which may furthermore be the element determining that the selected *root* is verbal, along lines discussed by Marantz (1997). Similarities between CP and DP suggest that DP too may be a phase, possibilities explored by Svenonius (2004) and Hiraiwa (2005) among others. I will put that aside here, and keep to the clausal skeleton – avoiding much structure here as well.

It seems problematic for T to fail to define a phase boundary along with C, since on the surface it seems to be T, not C, that is the locus of the φ-features that are involved in the Nominative-agreement system, and raising of the external argument subject or unaccusative/passive object to SPEC-T. There is, however, antecedent reason to suspect otherwise, confirmed (as we will see) by empirical phenomena. The antecedent reason is that for T, φ-features and Tense appear to be derivative, not inherent: basic tense and also tense-like properties (e.g., irrealis) are determined by C (in which they are inherent: "John left" is past tense whether or not it is embedded) or by the selecting V (also inherent) or perhaps even broader context. In the lexicon, T lacks these features. T manifests the basic tense features if and only if it is selected by C (default agreement aside); if not, it is a raising (or ECM) infinitival, lacking φ-features and basic tense. So it makes sense to assume that Agree- and Tense-features are inherited from C, the phase head.[26] If C-T agrees with the goal DP, the latter can remain in-situ under long-distance Agree, with all uninterpretable features valued; or it can raise as far as SPEC-T, at which point it is inactivated, with all features valued, and cannot raise further to SPEC-C.[27] We thus derive the A-A' distinction.

There is ample evidence that the A-A' distinction exists.[28] Let's assume it therefore to be a real property of I-language. The device of inheritance establishes the distinction in a simple way, perhaps the simplest way. The mechanism is a narrow violation of NTC. The usual question therefore arises: does it violate SMT? If it does, then the device belongs to UG (perhaps parametrized), lacking a principled explanation. But the crucial role it plays at the C-I interface suggests the usual direction to determine whether it is consistent with SMT though violating NTC. If the C-I interface requires this distinction,[29] then SMT will be satisfied

[25] Another case, brought up by a reviewer (citing Mel'chuk 1988 and Bošković 2004), is agreement into finite clauses in Chukchee.

[26] Sometimes the φ-features of C are morphologically expressed, as in the famous West Flemish examples. We leave open the question of how, or whether, expression of the features on C relates to the CP-internal syntax. See Miyagawa (in press) for review of a number of cases.

[27] In the framework of Chomsky (2001), "inactivation" of raised DP at SPEC-T was attributed to the fact that T is *complete*, with all features specified (unlike participles, lacking person). Those additional assumptions are unnecessary here; only completeness at the phase level has an effect. The earlier assumptions are also refuted by empirical phenomena in various languages, as discussed by Frampton and Guttman (2004) and Nevins (in press), who develop different approaches.

[28] See Rizzi (in press) for review, and a much more general context.

[29] To pursue the matter, one should ask such questions as whether quantifier-variable notation (as distinct from formally equivalent variable-free systems) is somehow an empirical property of the C-I (thought) systems, analogous to properties of the SM systems. Similar questions arise in other connections: e.g., is the familiar notation for sentential calculus, matching nested embedding in natural language, an empirical property of the C-I system, as distinct from Polish notation? Relative ease of learning and interpretation suggests a positive answer to both questions, which might turn out to bear on language design in interesting ways.

by an optimal device to establish it that violates NTC, and inheritance of features of C by the LI selected by C (namely T) may meet that condition. If so, the violation of NTC still satisfies SMT.

When φ-features appear morphologically at T without tense (or in participles, etc.), they should therefore be regarded as just a morphological effect of agreement, without significance in the syntactic computation.[30] The relative inability of TP to be moved or to appear in isolation without C gives some further reason to suspect that TP only has phase-like characteristics when selected by C, hence derivatively from C, though this is not a criterial property.[31] We turn later to further evidence for these assumptions.

Let's return first to the properties of EF. Suppose that EF permits free Merge to the edge, indefinitely. That yields a certain subcategory of recursive systems: with embedding, a pervasive feature of human language.[32] If an LI α enters a derivation and its EF is not satisfied, the resulting expression will often crash, but not always; say, the expression "No." If EF is minimally satisfied for α, then α has a complement, to which C-I will assign some interpretation; a theta role in some configurations. The predicate-internal subject hypothesis asserts that C-I will also assign a theta role to second EM to v*, that is, to its specifier, the external argument EA. If only phase heads trigger operations (as I will assume), then IM will satisfy EF only for phase heads; apparent exceptions, such as raising to SPEC-T, are derivative, via inheritance. Merge can apply freely, yielding expressions interpreted at the interface in many different kinds of ways. They are sometimes called "deviant," but that is only an informal notion. Expressions that are unintelligible at the SM interface may satisfy the most stringent C-I conditions, and conversely. And expressions that are "deviant" are not only often quite normal but even the best way to express some thought; metaphors, to take a standard example, or such evocative expressions as Veblen's "perform leisure." That includes even expressions that crash, often used as literary devices and in informal discourse, with a precise and felicitous interpretation at the interfaces. The only empirical requirement is that SM and C-I assign the interpretations that the expression actually has, including many varieties of "deviance."[33]

The label of an SO must be identifiable with minimal search, by some simple algorithm. Two obvious proposals, carried over from X-bar-theoretic approaches, are (2) and (3):

(2) In {H, α}, H an LI, H is the label

(3) If α is internally merged to β, forming {α, β} then the label of β is the label of {α, β}

These principles suffice for virtually every case.[34] Sometimes, however, they conflict. One example is the first step of every derivation, when two LIs merge, so that by (2), either may project. That seems

[30] See Iatridou 1988 on tenseless T with φ-features in Greek. See also Freidin (2004) for similar conclusions on different grounds.

[31] Thus TP cannot be clefted stranding C, etc. There are proposals to move TP in Kayne's LCA framework, to account for right-most complementizer and other phenomena. See also Rooryck (2000). These operations are unformulable in the present limited framework, at least for tensed CP, and even here do not separate TP from C, leaving TP adjacent to C and in its m-command domain – a notion that also departs from optimal assumptions about minimizing search.

[32] Note that we understand Merge(XP, YP) as satisfying EF of one or the other label – say the label of YP, if we think of XP as merged to YP – even if it is not an LI. Thus Merge of external argument to v*P satisfies the EF of v*. NTC requires that Merge in this case too is to the edge.

[33] These are the earliest assumptions of generative grammar, in Chomsky (1955) and later work. They still seem to me basically accurate, though the early proposals for assigning a status to expressions were quickly shown to be inadequate.

[34] The exceptions are EM of non-heads XP, YP, forming {XP, YP}, as in external argument merger of DP to v*P. The conventional assumption is that the label is v*. A possibility is that either label projects, but only v*-labeling will yield a

unproblematic, though one of the choices may yield some form of deviance. A more interesting case arises when LI α is internally merged to non-LI β. In this case, (2) yields the conclusion that α is the label, while (3) yields the conclusion that the label of β is the label. Consider, for example, *wh*-movement of the LI *what* to SPEC-C, forming (4), *t* a copy of *what*:

(4) what [C [you wrote *t*]]

If C projects, in accord with (3), then (4) can be, for example, the interrogative complement of "wonder" in "I wonder what you wrote." Iatridou et al. (2001) and Donati (in press) point out that *what* may also project, in accord with (2), yielding the free relative object of "I read [what you wrote]," interpreted as a DP headed by *what*. In conformity with (2), that is possible only when the phrase that is moved is a head; there can be no free relative interpretation in "I read [what book you wrote]." Relying on more subtle properties, Donati shows that the same is true for comparatives. Binding Condition (C) might be a case of (2) under EM, with the pronominal SPEC becoming the label, hence serving as the probe, as discussed earlier. Note that when *what* is taken to be the label in (4), so is C (for reasons of feature-inheritance by T). That is, the two labels coexist, in accord with a literal interpretation of the labeling algorithm (2)-(3).

The conclusion, then, is that the labeling algorithms apply freely, sometimes producing deviant expressions. The outcome will satisfy the empirical conditions on I-language if these are the interpretations actually assigned.

There must be some way to identify internally-merged α with its copy, but not with other items that have the same feature composition: to distinguish, say, "John killed John" or "John sold John to John" (with syntactically unrelated occurrences of *John*), from "John was killed John" (with two copies of the same LI *John*). That is straightforward, satisfying the inclusiveness condition, if within a phase each selection of an LI from the lexicon is a distinct item, so that all relevant identical items are copies. Nothing more than phase-level memory is required to identify these properties at the semantic interface C-I, where the information is required.

It has sometimes been suggested that IM should be eliminated in favor of EM. As noted, that necessitates a stipulation barring IM, and thus requires empirical support. It also requires some other device to distinguish copies from unrelated occurrences. And it involves additional computation when a phrase that would otherwise be internally merged (moved) has to be generated independently before merger, then somehow identified as a copy; perhaps generated many times, as in successive-cyclic movement. There are other rather severe complications resulting from the cyclic local character of IM, discussed below. I do not see any compensating advantages to this departure from SMT.

Since IM always leaves a copy, thanks to the NTC, and the copies are carried to the semantic interface, we eliminate the lowering operation of reconstruction – though just how reconstruction works is by no means a simple question. What about the phonetic interface? Here two desiderata conflict: (i) ease of processing, (ii) minimization of computation. Processing would be eased if all copies remain. That would eliminate many of the "filler-gap" problems that complicate parsing programs and perceptual theories. But minimization of computation calls for erasure of all but one copy, so that the phonological component can forget about the others; the issue does not arise in the mapping to the semantic interface, where all copies

---

coherent argument structure at C-I Another possible case is small clauses, if they are headless. A suggestive approach, along the general lines of Moro (2000), is that these structures lack a label and have an inherent instability, so that one of the two members of the small clause must raise.

remain without complication.[35] Overwhelmingly, (ii) is correct. That provides more evidence that language is "designed" so that mapping to C-I approximates the SMT, with utility for communication only a secondary factor.

If minimization of computation is the driving force in "Spell-Out" of copies, there should be an exception to the conclusion that only one is phonetically realized: namely, when special conditions, such as Lasnik's stranded affix filter, require some residue of the lowest copy to satisfy interface conditions. For evidence supporting that conclusion, see Abels (2001), Bošković (2001), Landau (2004), Hiraiwa (forthcoming).

For minimal computation, the probe should search the smallest domain to find the goal: its c-command domain. It follows that there should be no m-command, hence no SPEC-head relations, except for the special case where the SPEC itself can be a probe. That requires considerable rethinking of much important work, particularly on agreement. I think the conclusion is tenable,[36] but it is far from obvious. Without further stipulation, the number of specifiers is unlimited; the specifier-complement distinction itself reduces to first-Merge, second-Merge, etc. Again, these conclusions, restricting descriptive technology and thus approaching SMT, have to be shown to be empirically viable.

Minimal search is not uniquely defined in XP-YP structures where neither XP nor YP is a head: the "wrong choice" yields island effects. Among the means proposed to identify the right choice are Kayne's connectedness principle and government. Each involves stipulation of mechanisms that depart from SMT, and hence motivates a search for alternatives.

Consider the CED effects discovered by Huang (1982), involving XP-YP structures with island violations under the wrong choice. The adjunct-island subcase follows if an adjunct is not in the search domain of the probe. That in turn follows from the approach to adjuncts in Chomsky (2004), taking them to be entered into the derivation by pair-Merge instead of set-Merge to capture the fundamental asymmetry of adjunction,[37] then simplified to set-Merge at the point of Transfer, thus permitting phonetic linearization and yielding "late-insertion" effects at the semantic interface.

Consider the subject-island subcase. It has been assumed since Huang's discovery of these properties that the surface subject is the island, but there is reason to doubt this assumption. Compare (5) and (6):[38]

---

[35] Technically, this statement requires minor qualification: there is processing of each copy in the mapping to the semantic interface. But it is assumed (for good reasons) that these operations are universal, hence in effect instantaneous and costless, unlike the mapping to the phonetic interface, which tolerates substantial variety (again for obvious reasons) and therefore involves language-specific and sometimes complex computations.

[36] E.g., Reuland's observations on Condition (A), cited above. For another illustration, see McCloskey (2002), who argues that the form of the complementizer in Irish A'-binding does not depend on SPEC-head agreement, as had been supposed, but on how SPEC is formed – by EM or IM.

[37] For some similar ideas, see Åfarli (1995), Rubin (2003), Safir (1986). A slight modification might be needed, restricting simplification to the Transfer operations themselves, never entering narrow syntax. That depends on whether intervention effects and PIC will bar all other cases.

[38] Idealizing the judgments, and keeping to the pied-piping case, to exclude the incomplete constituent effects studied by Kuno (1972). In the oral tradition, including talks of mine, examples have kept to "picture-PP," but that lexical choice introduces extraneous issues because of the ambiguity of the phrase, which can be understood with PP interpreted not as a complement of "picture" but as, in effect, a reduced relative clause (roughly, "I have a picture which is of Boston," contrary to *"I saw a driver who is of the car," "I saw an author who is of the book"). The differences show up elsewhere, e.g., in *one*-replacement. Note that there are apparent counterparts to (5) and (6) in extraposition structures, but these diverge sharply, as we will see below, strengthening the conclusion already mentioned that PP-extraposition is not part of narrow syntax.

(5)(i) it was the CAR (not the TRUCK) of which [they found the (driver, picture)]

(ii) of which car did [they find the (driver, picture)?

(6) (i) *it was the CAR (not the TRUCK) of which [the (driver, picture) caused a scandal]

(ii) *of which car did [the (driver, picture) cause a scandal]

These are standard examples of the subject-island condition. The interesting case is (7):

(7)(i) it was the CAR (not the TRUCK) of which [the (driver, picture) was found]

(ii) of which car was [the (driver, picture) awarded a prize]

These fall together with (5), not (6), though the surface subject is in the same position as in (6). If so, then the effect is determined by the base structures of (7), not the surface structures, in which the distinction between the cases has been effaced by raising of the surface subject from the verb phrase. The relevant base structures are (8):

(8)(i) C [T [v [V [the (driver, picture) of which]]]]

(ii) C [T [$_{\alpha}$ [the (driver, picture) of which] [v* [V XP]]]]

In (i) v is unaccusative/passive, so that only (ii) but not (i) has the internal phase α. We now have the right distinction,[39] though it remains to explain it.

There are further consequences. One is that T is not the probe that yields A-movement of [*the* (*driver*, *picture*) *of which*] to the SPEC-T position in (7) before C is merged: if it were, the required distinction would again be effaced before *wh*-movement. Rather, A- as well as A'-movement must be triggered by probes in C: the probe for *wh*- accesses *which* in its base position in (7), raising *of-which* to SPEC-C, while the Agree-probe in C, inherited by T, raises the full DP [*the* (*driver*, *picture*) *of which*] to SPEC-T, the two operations proceeding in parallel. It follows further that TP is not a phase; rather CP, as already concluded on other grounds. Other considerations converge towards the same conclusion.

It remains to explain why the probe for *wh*-movement cannot readily access the *wh*-phrase within the external argument of α. That could reduce to a locality condition: *which* in α is embedded in the lower phase, which has already been passed in the derivation. We know that the external argument itself can be accessed in the next higher phase, but there is a cost to extracting something embedded in it – facts to which we return.[40]

Note that in (5), the problem does not arise. In (5), the *wh*-phrase is extracted to the edge of v*P unproblematically, then on to the edge of CP; nothing is extracted from it in the CP phase. In (6), the PP-complement of the subject cannot be extracted in the same way in the v*P phase, because its base position

[39] Choice of v* might have an effect. Perhaps "of which books did the authors receive the prize" is more acceptable than (6). If so, difference among theta roles might be relevant, perhaps requiring a deeper analysis of base structures.

[40] These considerations suggest a possible independent argument for the predicate-internal subject hypothesis. For some suggestions, on different grounds, about C as the locus of mood and agreement features, see Aygen (2004).

is not in the search domain of the label/probe v*; for the same reason, SPEC-to-SPEC movement is always impossible. Therefore extraction in (6) would have to be from the base position, distinguishing the cases properly.

We therefore reinforce the conclusion that C has two probes: the edge-feature EF that is automatically available for an LI, and an Agree-feature (φ-features). The former attracts the wh-phrase to the edge of C, the second attracts the DP, but only as far as T, with which it agrees. These facts raise the usual two questions: *how* (what are the mechanisms?) and *why* (what is the motivation?).

The obvious mechanism is the one already suggested for other reasons: T inherits its Agree-feature from C, and then derivatively serves as a probe at the phase level CP. The motivation may trace back to a C-I interface requirement that both arguments and operator-variable structures be available, analogous to the requirement of semantic duality that is satisfied in an optimal way by the A-A'-distinction, as already discussed.

On optimal assumptions, transmission of the Agree-feature should be a property of phase-heads in general, not just of C. Hence v* should transmit its Agree-feature to V, and probe of an object with structural Case by v* should be able to raise it to SPEC-V, a step-by-step analogue to raising to SPEC-T by C. That would yield the intriguing but puzzling conclusions about raising of objects to SPEC-V, particularly in ECM constructions, but perhaps generally.[41] The evidence is compelling, but has been unclear why such rules should exist: why should objects raise to SPEC-V at all, an operation that is even more odd because its effects are not visible, given that V then raises to v*? These strange facts fall into place automatically if the properties of the C-phase hold of phases generally. They thus yield further evidence to support the basic assumptions: C and v* are the phase heads, and their Agree-feature is inherited by the LI they select. Futhermore, if the suggestions above about motivation prove to be accurate, the curious phenomenon of raising to SPEC-V follows from the C-I requirement that the A-A' distinction must be observed at the CP-phase, supplemented by third-factor conditions of efficient computation.

As noted earlier, it is tempting to speculate that the resistance to extraction of the complement of C and v* (TP and VP respectively) is traceable to the requirement of feature inheritance from the selecting phase head. Note also that there is still an asymmetry between C-T and v*-V: T can appear without C, but V requires v*. That too would follow if lexical items are roots, with functional elements (v, n, etc.) determining their category, along lines already mentioned.

Another question is whether inheritance is obligatory or optional. For C-T, that raises familiar questions about universality of EPP and about mechanisms of agreement. For v*-V, properties of Binding Theory condition (B) indicate that the rule must be obligatory, by the general logic of the clause-mate principles of Postal-Lasnik-Saito. Thus in (9), *him* is necessarily free:

(9) the slave expected [(the picture, the owner) of him] to be somewhere else

The *for-to* analogue, to my ear, allows the bound option more readily, as would be expected. Lasnik, however, has given arguments to the contrary (see Lasnik 2002). The questions are unsettled for both C-T and v*-V.

[41] See Lasnik and Saito (1991); Lasnik (2003) for further discussion, and summary of history, back to Postal (1974).

What is true of (5)-(7) should hold in general for *wh*-questions. Consider the simpler cases (10), (11), with indices for expository purposes only:

(10)(a) C [T [who [v* [see John]]]]

(b) who$_i$ [C [who$_j$ [T [who$_k$ v* [see John]]]]]

(c) who saw John

(11) (a) C [T [v [arrive who]]]

(b) who$_i$ [C [who$_j$ [T [v [arrive who$_k$]]]]]

(c) who arrived

In (10), in the v*-phase v*-*John* agreement values all uninterpretable features. Turning to the C-phase, both the edge- and the Agree-feature of C seek the goal *who* in SPEC-v*. The Agree-feature, inherited by T from C, raises it to SPEC-T, while the edge-feature of C raises it to SPEC-C. The result is (10b). There is a direct relation between *who$_i$* and *who$_k$*, and between *who$_j$* and *who$_k$*, but none between *who$_i$* and *who$_j$*. There are two A-chains in (b): (*who$_j$, who$_k$*) and (*who$_k$*). Each A-chain is an argument, with *who$_i$* the operator ranging over the A-chains, interpreted as restricted bound variables.[42] Similarly in (11), with no lower phase, parallel operation of the edge- and Agree-features of C derives (b) from (a), with the operator *who* in SPEC-C and the two A-chain arguments (SPEC-T, Complement-V) and (Complement-V).

In both cases the A-chains are invisible, but familiar properties of A-movement (binding, scope, weak cross-over, etc.) reveal that there really is a copy in the position *who$_j$* heading the two-membered A-chain, even though it is not pronounced. Thus we have such standard distinctions as in (12):

(12)(a) who was never seen, *who was there never seen

(b) who$_i$ seems to his$_i$ friends to be preferable, *who$_i$ do you seem to his$_i$ friends to prefer

It has been conventionally assumed that in such constructions as (10), (11), there is an A-chain formed by A-movement of the *wh*-phrase to SPEC-T, and an A'-A chain formed by A'-movement of the subject to SPEC-C. There was never any real justification for assuming that there are two chains, a uniform A-chain and a non-uniform A'-A chain, rather than just one A'-A-A chain formed by successive-cyclic raising of the *wh*-phrase to SPEC-T and then on to SPEC-C. We now see, however, that the intuition is justified. There is no direct relation between the wh-phrase in SPEC-C and in SPEC-T, and no reason to suppose that there is a non-uniform chain at all: just the argument A-chains and an operator-argument/variable construction. By the usual demand of minimal computation, the A-chains contain no pronounced copy.

As matters now stand, constructions of the form (10), (11) can be formed in two ways: (I) as just described, with the edge feature of C extracting the wh-phrase from its base position, and (II) with the Agree feature

[42] Reformulation of application of operations can dispense with chains, if there is some reason to do this. I will keep to the chain notation for expository convenience.

of C-T forming the A-chain headed by SPEC-T (who_j in these examples), at which point the edge-feature EF of C raises who_j to SPEC-C.[43] (II) looks suspicious, both because it is redundant, and because we know that EF of C cannot extract the PP complement from within SPEC-T: if it could, the subject-condition effects would be obviated. It must be, then, that the SPEC-T position is impenetrable to EF, and a far more natural principle would be that it is simply invisible to EF, hence barring (II) as well. That principle generalizes the inactivity condition of earlier work, which takes the head of an A-chain (which always has any uninterpretable features valued) to be invisible to Agree.[44] A reasonable principle, then, is that an A-chain becomes invisible to further computation when its uninterpretable features are valued. That will incorporate the effects of the earlier inactivity condition, restated in terms of phases.[45]

We have been presupposing the informal notions A- and A'-position. For our purposes here, it will suffice to define an A'-position as one that is attracted by an edge-feature of a phase head; hence typically in SPEC-C or outer SPEC of v*. Others are A-positions. From this point of view, A- and A'-positions are distinguished not by their structural status within a phrase-marker, but by the manner in which they are derived. The shift of perspective has many consequences. I will tentatively assume that it is on the right track. It follows that successive-cyclic A'-movement creates a uniform A'-chain, no matter where the landing sites are along the way. Intermediate positions do not induce binding effects or have other A-position properties, whatever their structural status is.

Let us review some of the properties of the theory of IM. With all operations driven by the phase head, the only A-chains are completed A-chains with all features valued, either inherently or by Agree. It follows that "traces" (technically, lower copies) are invisible, as desired.[46] We also conclude that in A-movement, features are not valued until the operation is completed; otherwise the operation could not apply. Suppose PH is a phase head (C or v*) selecting PH_s. Then EF of PH can raise XP to SPEC-PH, but only from its base position. If the Agree feature of PH (inherited by PH_s) has raised XP to SPEC-PH_s, then XP is invisible; it cannot be raised and nothing can be extracted from it. If, however, EF raises XP to SPEC-PH, it no longer heads an A-chain (by definition), and is subject to raising or extraction by a higher EF. Extraction from this A'-position should be on a par with extraction from an external argument, carrying the cost of searching into a phase already passed. Thus (13) should have about the same status as the subject-island violations (6):[47]

(13) of which car did you wonder [which [(picture, driver)] [caused a scandal]]

Let's look more closely at legitimate application of EF of the phase head PH to XP in its base position, either raising XP to SPEC-PH, or raising its complement to SPEC-PH as in (5), (7). We know that raising cannot follow long-distance agreement valuing the features of XP, or XP would be invisible, for the reasons just given: its PP-complement could not be extracted, and XP itself could not be raised. Furthermore, XP cannot be raised before agreement, or its Case feature will be unvalued. It follows that

[43] As we will see directly, such interweaving of operations is permissible.
[44] If in SPEC-T, it is already invisible to IM, because of PIC.
[45] It remains the case that inherent Case will be visible to Match, inducing interference effects.
[46] That would also be true of the trace of a raised DP with quirky Case, adopting earlier assumptions that quirky Case involves a structural Case feature, valued by raising. From the formulation given here, it follows that the quirky-Case head of a non-trivial A-chain should not cause an interference effect. I do not know if it is possible to test this consequence.
[47] On such cases, see unpublished work by Esther Torrego, reviewed in Chomsky (1986), 26f.

the edge- and Agree-features of PH apply in parallel: EF raises XP or a PP complement within XP to SPEC-PH, while agreement values all uninterpretable features and may or may not raise XP to form an A-chain.[48]

We therefore have a rather delicate array of conclusions about the mechanisms of probe-goal relations and IM. They amount to the conclusion that all options are open: the edge- and Agree-features of the probe can apply in either order, or simultaneously, with only certain choices converging.

What holds for *wh*-movement should extend to A'-movement generally. Suppose that the edge-feature of the phase head is indiscriminate: it can seek any goal in its domain, with restrictions (e.g., about remnant movement, proper binding, etc.) determined by other factors.[49] Take, say, Topicalization of DP. EF of a phase head PH can seek any DP in the phase and raise it to SPEC-PH. There are no intervention effects, unless we assume that phrases that are to be topicalized have some special mark. That seems superfluous even if feasible, particularly if we adopt Rizzi's approach to the left periphery: what is raised is identified as a topic by the final position it reaches, and any extra specification is redundant. The same should be true for other forms of A'-movement. We need not postulate an uninterpretable feature that induces movement, and can thus overcome a long-standing problem about crash at the lower phase-levels in successive-cyclic movement.

Further elaboration depends on how the relevant structures are to be analyzed properly. To mention a few possibilities, suppose that the moved phrase is labeled by an interpretable interrogative *wh*-feature. Then it will have to reach the right position in the left periphery for interpretation, or be associated with such a position by some other operation.[50] Otherwise the expression may converge, but will be interpreted as deviant at the C-I interface. A *wh*-phrase lacking the interpretable interrogative feature, or an empty operator, will yield a structure that converges but will again have no interpretation unless the phrase undergoes A'-movement to the root, constituting a possible predicate or source for a head-raising relative.[51] Note that there should be no superiority effect for multiple *wh*-phrases; any can be targeted for movement. We are led to that conclusion for other reasons as well. Thus suppose that we have *what* in (10) instead of *John*:

(14) C [T [who [v* [see what]]]]

At the lower v* phase the subject *who* does not intervene, so even if uninterpretable *wh*-features are targeted by the edge-feature of v*, *what* can be raised to the edge, voiding any superiority effect. That leaves the problem of explaining the superiority phenomena in the languages in which they appear: English, but apparently not German in simple cases, for example. Standard examples, such as (14), tell us very little, and it is not so simple to find convincing cases.[52]

---

[48] There is nothing problematic about application of features in parallel. It has always been assumed, unproblematically, for probing by φ-features.

[49] That should be the case for independent reasons, since EF-probe does not involve feature matching, hence Agree.

[50] See among others Pesetsky (1987), Fiengo et al. (1988), Reinhart (1993), Tsai (1994), on options for *wh*-in-situ languages. For extensive discussion of many related issues, see Cheng and Corver (in press).

[51] For evidence that both kinds of relatives can co-exist, head-raising and head-Merge, see Szczgielniak (2004).

[52] See Chomsky (1995), chap. 4, note 69. The basic problem with the standard cases is that the in-situ *wh-phrase* is the prosodic peak, and might have wide scope under a focus interpretation. The problem is compounded by the fact that the pair-list interpretation disappears if stress is shifted, as in "who NEVER saw what"? See Nissenbaum (2000), chap. 3, note 2, for observations on deaccenting, which can be stated in terms of superiority, though it remains to explain them. Standard efforts to

Do A'-chains function in the manner of A-chains with regard to intervention? The line of argument we are pursuing indicates that they should. That appears to be correct. Consider such constructions as "who did John see," schematically (15):

(15) C [T [John [v* [V who]]]]

Given PIC, at the lower v* phase *who* raises to the edge, so it is accessible to the phase head C. The Agree-feature of C-T seeks the subject *John* and raises it to SPEC-T, and the edge-feature of C seeks the object *who* in the outer SPEC of v* and raises it to SPEC-C. If T were a phase head, or an independent probe for some other reason (as assumed in earlier work, mine in particular), then raising of subject to SPEC-T would be blocked by intervention of the φ-features of *who* in the outer SPEC of v*. But since it is not a phase head, and both operations are driven by the phase head C in parallel, the problem does not arise. However, raising of the subject does cross the lower copy of *who* in the A'-position of the outer SPEC of v*, that is, the lower copy in the A'-chain (SPEC-C, OUTER-SPEC-v*). A'-chains thus behave in this respect like A-chains: if uniform, only the full chain (equivalently, its head) is the object that intervenes. We therefore have uniform chains -- either A-chains or A'-chains – but no mixed chains.[53]

A somewhat more complex illustration similar to (5)-(7) has been discovered in Icelandic.[54] Keeping just to the essence, consider the dative-nominative experiencer construction (16) (e.g., "to-someone seems [the horses are slow]"):

(16) C [T [DAT [v* NOM...]]]

If DAT remains in-situ, in an expletive construction, it blocks T-NOM agreement, as expected. If DAT is raised to SPEC-T, T-NOM agreement is permitted, again as expected: there is no intervening argument. But if DAT is *wh*-moved, it blocks agreement, which is paradoxical since it appears to be the lower copy of an A-chain. The solution suggested by Holmberg-Hroarsdottir (along with several other devices) is that the DAT subject moves only to SPEC-C, directly, so that the sole A-chain has only one position, its base position, which blocks agreement by intervention. That seems basically right, but it yields new problems, because we do want to have the A-chain (SPEC-T, SPEC-v*), for reasons already mentioned. The desired result follows, as before, if both operations, A- and A'-movement, are driven by the phase head C. That will leave no relation between SPEC-C and SPEC-T, but an operator-argument relation between SPEC-C and each of the two A-chains, (SPEC-T, SPEC-v*) (formed by the Agree-feature of C) and (SPEC-v*) (formed by the edge-feature of C). The A-chain (SPEC-v*) suffices to block T-NOM agreement, by standard intervention; the A-chain (SPEC-T, SPEC-v*) yields the A-movement effects.

One might explore whether the variation in judgments (see note 50) can be attributed to the timing of the edge-and Agree probes.

---

account for superiority in terms of locality do not apply, at least in any obvious way. See Pesetsky (2000) on what may be superiority effects in more complex constructions.

[53] Notice again that we are again glossing over the question of the precise mechanisms, which can be formulated in various ways, including elimination of chains altogether. The questions are interesting, but do not seem to bear directly on this level of discussion.

[54] Hiraiwa (2002, 2005); Holmberg and Hroarsdottir (2003), who observe that not all speakers accept the data they describe.

The examples discussed so far are restricted to two-membered A-chains, as in (5-7), (10-11), (16). Consider now successive-cyclic A-movement.[55] Compare (17) = (6) with (18), in both cases with bold-face brackets around the internal phases and the other brackets around TP, and *t* used for the lower copies in the A-chains:

(17) (i) *it was the CAR (not the TRUCK) of which [the (driver, picture) **[***t* caused a scandal**]**]

(ii) *of which car did [the (driver, picture) **[***t* cause a scandal**]**]

(18) (i) it is the CAR (not the TRUCK) of which [the (driver, picture) is likely [*t* to **[***t* cause a scandal**]**]]
(ii) of which car is [the (driver, picture) likely [*t* to **[***t* cause a scandal**]**]]

In (17), as already discussed, the PP phrase in the EA "the (driver, picture) of which" raises to SPEC-C from its base position, and we have the subject-island effect. But in (18), the effect is obviated. These expressions have the status of extraction from object, not subject, as in (5), (7). That follows from the previous conclusions about IM, assuming the (still mysterious) condition EPP, which requires A-movement to pass through SPEC-T, with familiar consequences for binding, and possibly reconstruction. One permitted order of operations is this: the Agree-feature of C-T raises EA step-by-step to its final position, and along the way, the edge-feature of C extracts the PP complement and raises it to SPEC-C, with no deep search required because no phase boundaries are crossed. The parallel operations interweave, again unproblematically.

Note that the same conclusions hold, as expected, for ECM. Thus (19) has the status of (18), not (17):

(19) of which car did they believe the (driver, picture) to have caused a scandal

It must be, then, that "of which car" is raised from an intermediate position, SPEC-T of the ECM infinitival, before it reaches SPEC-V, a position analogous to SPEC-T in the matrix clause. In SPEC-T and SPEC-V, all features are valued in the completed A-chain, and its head is invisible, as we have seen.

Reinforcing earlier conclusions, we find that the two searches driven by the phase head operate in parallel, and can even interweave. What yields the subject-island effect, it appears, is search that goes too deeply into a phase already passed, not the difference between base and surface position. The Agree-feature of C raises XP to SPEC-T (by inheritance), while its edge feature raises XP (or part of it) to SPEC-C. The generalized inactivity condition bars extraction from matrix SPEC-T (and if fully generalized, raising of full XP from that position). Such constructions as (19) provide an independent reason for ECM-raising to SPEC-V. And extraction of complement provides an independent reason, alongside of binding and (perhaps) reconstruction effects, for successive-cyclic A-movement through SPEC-T.

Note that raising of the PP-complement is sharply different from extraposition of PP in such constructions. Thus (20) has the same status as extraposition from EA, as in (21), and both are much worse than the (relatively weaker) subject-island effect of (17):

(20) *the (driver, picture) is likely to cause a scandal of the car

[55] Cases brought up by Sam Epstein, pc.

(21) *the (driver, picture) caused a scandal of the car

We conclude, then, that despite some superficial similarities, extraposition and A'-movement of PP-complement are entirely different phenomena. Quite possibly, as suggested earlier, PP-extraposition is part of the mapping to the SM interface, hence is part of SPELL-OUT. If so, it should be restricted to the interior of a phase, thus allowing extraposition from object but not from EA subject.[56]

As discussed elsewhere (Chomsky 2001), the size of phases is in part determined by uninterpretable features. Such features are a striking phenomenon of language that was not recognized to be significant, or even particularly noticed, prior to Vergnaud's original ideas about the role of structural Case (see note 1). The values of these features are redundant, fixed by structural position in the course of derivation. We therefore expect them to be selected from the lexicon unvalued. Since these features have no semantic interpretation, they must be deleted before they reach the semantic interface for the derivation to converge. They must therefore be deleted either before Transfer or as part of Transfer. But these features may have phonetic realization, so they cannot be deleted before transfer to the phonological component. They must therefore be valued at the stage in computation where they are transferred – by definition, at the phase-level – and this must be the same stage for both transfer operations, again supporting the optimal assumption that transfer to both interfaces is at the same stage of derivation. Once valued, the uninterpretable features are deleted by the mapping to the semantic component, and given whatever phonetic properties they have in particular I-languages by the phonological component. The conclusion is supported by the fact that once features are valued, they are indistinguishable from interpretable features and there is no indication of their relation to the interpretable features that match them and assign them their values.[57] Hence they must be transferred at the point where they are valued: again, at the phase level, assuming that all operations apply at this level, as determined by the label.

These observations provide further support for the conclusion that v*P and CP are phases, the locus of determination of structural Case and agreement for object and subject. For subject, the conclusion is based on the assumption that TP is not a phase, for reasons discussed, so that T operates as a probe only derivatively by virtue of its relation to C (similarly, v* and V). A stronger principle would be that phases are exactly the domains in which uninterpretable features are valued, as seems plausible.

There is also morphological evidence that CP, v*P are the phases. Just for these two categories the edge is sometimes morphologically marked in successive-cyclic movement, with the effect of movement through SPEC-C sometimes found in the subject-agreement domain, another reason to suspect that T-agreement is derivative from properties of C.[58] These also seem to be the stages that permit parasitic gap constructions; for extensive discussion, based on the assumption that these are phases for A'-movement, see Nissenbaum (2000). There is also evidence for other effects of phases, among them in phonology (Ishihara, 2003) and covert movement (Cecchetto, 2004), and there is also much further investigation of alternatives and further articulation that I cannot review here.

Phases should, presumably, be as small as possible, to minimize computation after Transfer and to capture as fully as possible the cyclic/compositional character of mappings to the interface. C and v* impose an upper bound, and T is too small, as we have seen. More generally, there are two basic cases to consider:

[56] Something similar appears to be true with regard to peculiarities of in-situ object in English; see Chomsky (2001).
[57] Assuming no resort to extra mechanisms, which can always be devised, but require independent justification.
[58] See sources cited in Chomsky (2004), and for further elaboration, several papers in Cheng and Corver (in press).

(I) SO cannot be transferred to the SM interface ("spelled out") if it is subsequently going to move

(II) SO cannot be moved to an edge unless it can be spelled out right there, satisfying any uninterpretable features by long-distance Agree

(I) is transparent, unless more complex apparatus is introduced that we would hope to avoid. (II) has to be sharpened. It conforms to a fairly general empirical observation that should be captured:

(III) In a probe-goal relation, the goal can be spelled out only in-situ (under long-distance Agree) or at the probe (under internal Merge)

Standard illustrations are passive/unaccusative in-situ or with movement. The goal cannot stop at some intermediate point of the derivation, in particular, at intermediate SPEC-T positions through which it must pass in successive-cyclic A-movement (including ECM constructions).[59] In the case of A'-movement, reconstruction effects indicate that the raised goal also passes through internal positions leaving copies that are visible at the semantic interface.[60] These observations tell us something important about the operation of IM: the raised goal must reach the probe by means of local steps, passing through intermediate positions where it leaves copies, but not stopping there to be spelled out.[61] Just how small these local steps are remains to be clarified. For A'-movement, they could turn out to be as small as every category, as proposed more generally on completely different (and long abandoned) grounds in Chomsky (1986). As noted earlier, there is an exception, which may have principled grounds: inheritance of the features of the phase-head by the category it selects, T and V. These properties of IM, which appear to be quite general, add further reason to suppose that operations are at the phase level only, and that inactivation of SPEC-T in a tensed clause is a reflex of inheritance of C features.

There are some asymmetries between A- and A'-movement with regard to local steps. One is that the reconstruction effects are far weaker for A-movement (if they exist at all[62]). The only strong argument for local steps for A-movement are those based on binding and (as discussed above) extraction. In the latter case, the argument supports only the option, but not the necessity, of the local step. In both cases the effects hold only at SPEC-T, hence fall within the EPP category. Furthermore, there is strong evidence that raising of EA to SPEC-T does not pass through intermediate positions (hence presumably that A-movement never does). If there were intermediate positions between the base and surface position in this case (say, at the edge of a participial phrase), then subject-island effects would be obviated, exactly as they are in successive-cyclic (and ECM) raising.[63] These properties remain to be explained.

A consequence of the conclusion that the Agree-feature belongs to C, and to T only derivatively, is that it is in the same region as the left-periphery head for Focus. This conclusion is developed by Miyagawa (in

[59] A long-standing problem, for which a number of devices have been proposed, none really satisfactory. An apparent exception in Icelandic might be related to the fact that transitive expletives are allowed, as discussed elsewhere. For similar conclusions, see Bošković (2002).
[60] On interpretive effects, see Legate (2003), and more generally, Fox (2000). On morphologically visible agreement at stages intermediate between probe-goal (whether the goal remains in situ or moves to the probe), see Boeckx (2004a).
[61] I am adapting here proposals of Boeckx (2003), based on work by Takahashi (1994), and earlier ideas phrased in terms of a minimal link condition.
[62] See Lasnik (2003), chaps. 9, 10; the issue of Freidin-Lebeaux effects that he discusses is not relevant here.
[63] John Frampton, pc.

press) to argue that agreement and Focus are two values of the same parameter, with languages varying as to which of them is prominent: φ-features for English-type languages, Focus for Japanese-type languages (including Bantu and others discussed by Baker (2003), whose proposals he adapts). Strengthening ideas on universality of features in Sigurdsson (2004), Miyagawa argues that these functional features are not only present in all languages, but are also phonologically expressed in some fashion. The relevant point here is that if analysis along these lines can be sustained, it provides further indirect evidence to support the conclusion that the phases are CP and v*P, the locus of valuation of uninterpretable features, placing Vergnaud's seminal ideas of 25 years ago in a much broader context and carrying the MP some important steps forward.

Let us turn finally to the mysterious property EPP, which has been an annoying problem ever since it was originally formulated to describe the obligatory presence of expletives in subject position of English-type languages if nothing raises to that position. EPP problems are considerably more general, however. Thus, while v typically permits both long-distance agreement and raising, v* does not:[64]

(22) *there will [a student [v* [take the class]]]

Rather generally, it seems, if languages of the relevant typology lack an expletive, the closest noun phrase raises to SPEC-T, sometimes with default agreement morphology (or none) on T.[65]

For infinitivals, we may be able to disregard control structures: subject is null so its structural position is uncertain, and they presumably fall under CP structures in any event. That leaves raising/ECM infinitivals. For these, EPP may in part be reducible to the general step-by-step property of IM, already discussed; but only in part, because of the special role of SPEC-T in such operations, a residual EPP effect. It also remains to account for EPP in tensed clauses (the C-T category) and the analogous v*-V issues.

It is tempting to ask whether EPP can be reformulated in terms of feature inheritance. Suppose that EF can be inherited from the phase head along with the Agree-feature. Not being a phase head, T need have no option for second-Merge by IM, but rather inherits it from C, and by some kind of feature-spread, this extends to all T's in the phase.[66] Operations then proceed as before. If there is no accessible NOM, then T will have default morphology, as in Icelandic and the Slavic constructions discussed by Lavine and Freidin; or null morphology, as in Miyagawa's Japanese examples. And there are a few other options.[67] If nothing is raised, then the inherited edge feature of T must be satisfied by EM, necessarily of an expletive since no argument role can be assigned. Possibly reformulation of the EPP properties in such terms might open a way to resolving the problems they raise. If so, it would be a welcome development, another step towards the goals of the MP and the long tradition of inquiry from which it derives.

---

[64] It has been argued that this property reduces to the requirement that something must evacuate v*P (Alexiadou and Anagnostopoulou, 2001). On subject in-situ in Icelandic transitive expletive constructions, see Jonas (1995, 1996).

[65] See, among others, Lavine and Freidin (2002), Miyagawa (in press). In the present framework, the Lavine-Freidin account has to be slightly revised so that it is C, not v, that assigns the (unexplained) structural accusative case in object position of unaccusative/passive in Slavic.

[66] It does not matter whether Agree also spreads, given the revised version of the inactivity condition.

[67] An implicit assumption here is that locative inversion in English (e.g., "in the square stood a statue") is to the same position as in sentences with overt subjects ("in the square, there stood a statue"), perhaps with null expletive. That seems to account for the clearest cases and their restrictions (e.g, the lack of V-C movement as in *"did in the square stand a statue"), though problems remain. On languages with clear locative-inversion of DP to subject, see Baker (2003) and Miyagawa (in press).

*To appear in C. P. Otero et. al. eds., *Foundational Issues in Linguistic Theory* (MIT)

Bibliography

Abels, K. (2001). The predicate cleft constructions in Russian. S. Franks, ed., *Proceedings of FASL* 9.

Åfarli, T. (1995) A note on the syntax of adverbial phrases. *Norsk Lingvistisk Tidsskrift* 13: 23-40.

Alexiadou, A., and E. Anagnostopolou (2001). The Subject-in-Situ Generalization and the Role of Case in Driving Computations. *Linguistic Inquiry* 32, 193-231.

Aygen, G. (2004). *Finiteness, Case and Clausal Architecture*. *MIT Occasional Papers*, v 23; 2002 Harvard Phd dissertation.

Baker, M. (2003). Agreement, dislocation, and partial configurationality. In A. Carnie, H. Harley, and M. Willie (eds.), *Formal Approaches to Function in Grammar* (John Benjamins).

Belletti, A., ed. (2004). *Structures and Beyond – The Cartography of Syntactic Structure*, Vol 3 (Oxford).

Boeckx, C. (2003). *Islands and Chains: Resumption as Stranding* (John Benjamins).

Boeckx, C. (2004a). Long-distance agreement in Hindi: some theoretical implications. *Studia Linguistica* 58 (1): 1-14.

Boeckx, C. (2004b). *Bare Syntax*. Ms, Harvard.

Boškovič, Ž. (2001). *On the Nature of the Syntax-Phonology Interface* (Elsevier).

Boškovič, Z. (2002). A-movement and the EPP. *Syntax* 5: 167-218.

Boškovič, Z. (2004). On the locality of Move and Agree: Eliminating the Activation Condition, Generalized EPP, Inverse Case Filter, and the Phase-Impenetrability Condition. Ms., University of Connecticut. (to appear in *U. Conn Working Papers Occasional Papers in Linguistics*).

Brody, M. (1995) *Lexico-Logical Form: A Radically Minimalist Theory* (MIT).

Cecchetto, C. (2004). Explaining the locality conditions of QR: Consequences for the theory of phases. Ms., Milan.

Cheng, L. and N. Corver, eds. (in press). *Wh-movement on the Move* (MIT).

Chomsky, N. (1955). *Logical Structure of Linguistic Theory*, ms. Parts of revised 1956 version published in 1975 (Plenum, Chicago).

Chomsky, N. (1965). *Aspects of the Theory of Syntax* (MIT).

Chomsky, N. (1986). *Barriers* (MIT).

Chomsky, N. (1993). A Minimalist Program for Linguistic Theory. In K. Hale and S.J. Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger* (MIT).

Chomsky, N. (1995). *The Minimalist Program* (MIT).

Chomsky, N. (2001). Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: a Life in Language* (MIT).

Chomsky, N. (2004). Beyond Explanatory Adequacy. In Belletti (ed.).

Chomsky, N. (2005). Three Factors in the Language Design. *Linguistic Inquiry* 36: 1-22.

Chomsky, N., and H. Lasnik (1993). The Theory of Principles and Parameters, in J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research* (de Gruyter). Reprinted in Chomsky (1995).

Cinque, G. (1999) *Adverbs and Functional Heads* (Oxford).

Cinque, G., ed. (2002). *Functional Structure in DP and IP – The Cartography of Syntactic Structures*, Vol. 1 (Oxford).

Collins, C. (2002). Eliminating Labels. In Epstein and Seely (2002).

Donati, C. (in press). On *wh*-head-movement. In Cheng and Corver (in press).

Epstein, S., and D. Seely (2002), *Derivation and Explanation in the Minimalist Program* (Blackwell).

Fiengo, R., C-T. J. Huang, H. Lasnik, and T. Reinhart (1988). The Syntax of *wh*-in-Situ. *Proceedings of West Coast Conference on Formal Linguistics* VII, 81-98.

Fox, D. (2000). *Economy and Semantic Interpretation* (MIT).

Frampton, J. and S. Guttman (2000). Agreement as Feature Sharing. Ms., Northeastern.

Frampton, J., and S. Guttman (2004). How Sentences Grow in the Mind. Ms., Northeastern.

Frampton, J., S. Guttman, C. Yang, and J. Legate (2000). Some Inconsistencies in the DBP Feature Valuation Algorithm. Ms., Northeastern.

Freidin, R. (2004) *Syntactic Structures* Redux. *Syntax* 7: 100-126.

Freidin, R., and J.-R. Vergnaud (2001). Exquisite Connections: Some Remarks on the Evolution of Linguistic Theory. *Lingua* 111: 639-666.

Hiraiwa, K. (2002). Multiple Agree. 25th GLOW-TiLT workshop, Utrecht.

Hiraiwa, K. (2005). *Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture*. MIT Phd dissertation.

Hiraiwa, K. (forthcoming). Dimensions of Agreement. To appear in C. Boeckx (ed.), *Complex Agreement Systems* (Oxford).

Holmberg, A., and T. Hroarsdottir (2003). Agreement and Movement in Icelandic Raising Constructions. *Lingua* 113: 997-1019.

Huang, C-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. MIT Phd dissertation.

Iatridou, S. (1988). On Nominative Case Assignment and a few Related Things. Ms, MIT. *MIT Working Papers in Linguistics 19: 175-196*, *Papers on Case and Agreement II* (1993).

Iatridou, S., E. Agagnostopoulos, and R. Izvorksi. Observations about the Form and Meaning of the Perfect. In M. Kenstowicz (ed.), *Ken Hale: a Life in Language* (MIT).

Ishihara, S. 2003. *Intonation and Interface Conditions*. MIT Phd dissertation.

Jonas, D. (1995). *Clause Structure and Verbal Syntax in Scandinavian and English*. Harvard Phd dissertation.

Jonas, D. (1996). Clause Structure, Expletives, and Verb Movement. In W. Abraham, S.D. Epstein, H. Thrainsson and C. J-W. Zwart (eds.), *Minimal Ideas* (John Benjamins).

Joos, M., ed. (1957), *Readings in Linguistics* (Washington: American Council of Learned Societies)

Kayne, R. (1981). Unambiguous Paths. In R. May and J. Koster (eds). *Levels of Syntactic Representation* (Reidel).

Kuno, S. (1972). Functional Sentence Perspective. *Linguistic Inquiry* 3:269-320.

Landau, I. (2004). Chain Resolution in Hebrew V(P)-fronting. Ms., Ben-Gurion.

Lasnik, H. (2002). Clause-mate conditions revisited. *Glot International* 6(4), April.

Lasnik, H. (2003). *Minimalist Investigations in Linguistic Theory* (Routledge).

Lasnik, H. (in press). Cyclicity. To appear in Cheng and Corver.

Lasnik. H. and M. Saito (1991). On the Subject of Infinitives. In L. Dobrin et al. (eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*. Reprinted in H. Lasnik, *Minimalist Analysis* (Blackwell 1999).

Lavine, J., and R. Freidin (2002). The Subject of Defective T(ense) in Slavic. *Journal of Slavic Linguistics* 10: 253–89.

Legate, J. (2003). Some Interface Properties of the Phase. *Linguistic Inquiry* 34: 506-16.

Lewontin, R. (1990). The Evolution of Cognition. In D. Osherson and E. Smith, eds., *An Invitation to Cognitive Science: Thinking*, vol. 3 (MIT).

Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. A. Dimitriadis et al. (eds.), *Proceedings of the 21st Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics* 4:2.

McCloskey, J. (2002). Resumption, Successive Cyclicity, and the Locality of Operations. In Epstein and Seely.

McGinnis, M. and N. Richards, eds. (in press). *Proceedings of the EPP/Phase Workshop*, *MIT Working Papers in Linguistics.*

Mel'cuk, I (1988). *Dependency syntax: Theory and practice* (SUNY).

Miller, G., and N. Chomsky (1963). Finitary Models of Language Users. In R.D. Luce, R. Bush, and E. Galanter, eds., *Handbook of Mathematical Psychology*, vol. II, chap. 13 (Wiley).

Miyagawa, S. (in press). On the EPP. In McGinnis and Richards.

Moro, A. (2000). *Dynamic Antisymmetry: Movement as a Symmetry Breaking Phenomenon* (MIT).

Nevins, A. (in press). Derivations without the Activity Condition. In McGinnis and Richards.

Nissenbaum, J. (2000). *Investigations of Covert Phrase Movement*. MIT Phd dissertation.

O'Neil, W. (1977). Clause Adjunction in Old English. *General Linguistics* 17.4 199-211.

Pesetsky, D. (1987). WH-in-situ: Movement and Unselective Binding. In E. Reuland and A.G.B. terMeulen, eds., *The Linguistic Representation of (In)definiteness*, (MIT).

Pesetsky, D. (2000). *Phrasal Movement and Its Kin* (MIT).

Postal, P. (1974). *On Raising* (MIT).

Reinhart, T. (1979). The Syntactic Domain of Syntactic Rules. In F. Guenther and S. Schmidt (eds.), *Formal Semantics and Pragmatics* (Reidel).

Reinhart, T. (1993). Wh-In-Situ in the Framework of the Minimalist Program." Ms., Tel Aviv. *Natural Language Semantics* 6: 29-56, 1997.

Reuland, E.J. (2001). Primitives of Binding. *Linguistic Inquiry* 32.3, 439-492.

Richards, N. (2001). *Movement in Language* (Oxford).

Rizzi, L. (1997). The Fine Structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar* (Kluwer). Reprinted in L. Rizzi, *Comparative Syntax and Language Acquisition* (Routledge, 2000).

Rizzi. L., ed. (2004). *The Structure of CP and IP – The Cartography of Syntactic Structures*, Vol. 2 (Oxford).

Rizzi, L. (in press). On the Form of Chains: Criterial Positions and ECP Effects. To appear in Cheng and Corver.

Rooryck, J. (2000). *Configurations of Sentential Complementation* (Routledge).
Rubin, E. (2003) Determining Pair-Merge. *Linguistic Inquiry* 34.4 660-68.

Safir, K. (1986). Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17: 663-689.

Sigurdsson, H. (2004). Meaningful Silence, Meaningless Sounds. Ms., Lund.

Svenonius, P. (2004). On the Edge. In D. Adger, C. de Cat and G. Tsoulas (eds.), *Peripheries: Syntactic Edges and their Effects* (Kluwer).

Szczegielniak, A. (2004). *Relativization and Ellipsis*. Harvard Phd dissertation. *MIT Occasional Papers in Linguistics*, Number 24, 2005. Cambridge MA.

Takahashi, D. (1994). *Minimality of Movement*. U. of Connecticut Phd dissertation.

Tsai, W.-T. D. (1994). *On Economizing the Theory of A-bar Dependencies*. MIT Phd dissertation.

Vergnaud, J.-R. (1985). *Dépendance et niveaux de représentation en syntaxe. Lingvisticae Investigationes Supplementa* 13, 1985 (John Benjamins).

Wexler, K., and P. Culicover (1980). *Formal Principles of Language Acquisition* (MIT).

Yang, C. (2002). *Knowledge and Learning in Natural Language* (Oxford).