

Foundations of Minimalism  
Howard Lasnik  
University of Maryland  
lasnik@umd.edu

Minimalism is a research program based on the following interrelated questions:

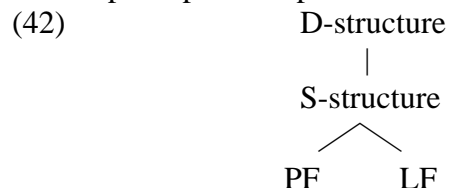
- (1) a. To what extent is human language a "perfect" system?  
b. To what extent is the computational system for human language optimal?
- (2) A human language is fundamentally a system for generating an infinite set of interpretable expressions. MP hypothesizes that the core operation necessary for this, Merge, is the only syntactic operation there is.
- (3) Take a lexical item to be a syntactic entity. Also take the result of combining a syntactic entity with another to again be a syntactic entity. Merge is the operation that effects the combining. Merge is reminiscent of the generalized transformations (GTs) of early transformational generative grammar (Chomsky (1955), LSLT), but extended in two respects:
- (4) First, it is responsible for **all** structure building, whereas in LSLT phrase structure rules built single clausal structures, with GTs combining these into complex structures, typically by embedding one in another.
- (5) Second, Merge subsumes the effects of the singularly transformations (movement and the like) of LSLT. This instance of Merge is 'Internal Merge', merging a syntactic entity that is already in a structure with that structure. [External Merge is the merging of something not already in a structure with that structure.]
- (6) Consider the derivation of "The woman will see the man":
- (7) The noun (N) *man* is combined with the determiner (D) *the* to form the determiner phrase (DP) *the man*.
- (8) This DP then combines with the verb *see* to produce an intermediate projection of V.
- (9) The DP *the woman* is created in the same fashion as *the man*, and is combined with the projection of V to produce the full VP.
- (10) Next, this VP merges with the Infl *will* producing a projection of I.
- (11) The DP *the woman* finally moves to the specifier position of I (i.e., internally merges with I), yielding the full clausal projection IP, as shown in (12):
- (12) [<sub>IP</sub> [<sub>DP</sub> The woman] [<sub>I</sub> will [<sub>VP</sub> t [<sub>V'</sub> see [<sub>DP</sub> the man]]]]]
- (13) Under minimalism, the traditional labels on the projections cannot actually arise because of Inclusiveness, nor can 'trace', a new syntactic entity be created:
- (14) Inclusiveness  
"any structure formed by the computation (in particular, PF and LF) is constituted of elements already present in the lexical items selected for [the numeration, the initial selection of lexical items] N; no new objects are added in the course of computation apart from rearrangements of lexical properties (in particular, no indices, bar-levels in the sense of X-bar theory, etc.)." Chomsky (1995, p. 228)

- (15) [<sub>I</sub> [<sub>the</sub> The woman] [<sub>I</sub> will [<sub>see</sub> [<sub>the</sub> The woman] [<sub>see</sub> see [<sub>the</sub> the man]]]]]]
- (16) With Merge as the operation that constructs phrase structure and also accounts for displacement phenomena ('movement'), it is technically impossible to construct canonical D-structure, which in previous theory is a level of representation that results from the application of phrase structure rules plus lexical insertion but crucially no movement transformations.
- (17) This technical argument meshes well with Chomsky's initial conceptual argument for LF and PF as the only levels of representation, hence against D-structure (and S-structure):
- (18) "[UG] must specify the interface levels (A-P, C-I), the elements that constitute these levels, and the computations by which they are constructed. A particularly simple design for language would take the (conceptually necessary) interface levels to be the only levels. That assumption will be part of the 'minimalist' program I would like to explore here." Chomsky (1993, p. 169)
- (19) This elimination of D-structure is, in interesting respects, a return to the roots of transformational generative grammar.
- (20) In the earliest work in transformational grammar (Chomsky (1955)) there is no level of D-structure.
- (21) As noted above, a phrase structure component constructs phrase-markers for simple sentences. Generalized transformations combine these single clause structures into multiple clause structures. The 'recursive' component of the grammar is thus the transformational component, in particular the generalized transformations, which combined pairs of P-markers either by coordinating the pair or by subordinating one to the other through embedding.
- (22) Chomsky (1965) rejects this model in favor of one with recursion in the phrase structure rule component. The output of this component and the lexical insertion transformations is 'deep structure'.
- (23) Chomsky's major arguments for this innovation are that it results in a simpler overall theory, and at the same time it explains the absence of certain kinds of derivations for which there appeared to be no empirical motivation.
- (24) Chomsky's second point is based on the observation in Fillmore (1963) that while there is extensive ordering among singular transformations, "... there are no known cases of ordering among generalized transformations although such ordering is permitted by the theory of Transformation-markers" (Chomsky 1965:133). Further, while there are many cases of singular transformations that must apply to a constituent sentence before it is embedded, or that must apply to a 'matrix' sentence after another sentence is embedded in it, "... there are no really convincing cases of singular transformations that must apply to a matrix sentence before a sentence transform is embedded in it..." (1965:133). In other words, the earlier theory allows for a class of grammars that do not appear to exist.
- (25) As for the argument from simplicity, Chomsky claimed that the theory of transformational grammar is simplified by this change, since the grammatical machinery of "generalized transformations" and "Transformation-markers" (T-markers) are eliminated entirely. The P-markers in the revised theory, what Chomsky designates as "generalized P-markers", contain all of the information of those in the LSLT version, but

- they also indicate explicitly how the clauses are embedded in one another, information that had been provided by the embedding transformations and T-markers.
- (26) This extension of the theory of phrase structure rules to include recursion, which makes generalized transformations redundant, also has consequences for the theory of singularly transformations.
  - (27) As indicated above, in the Aspects theory, as in the LSLT theory, there is extensive ordering among singularly transformations. In both frameworks, the set of singularly transformations was seen as a linear sequence: an ordered list. Given the Aspects modification, this list of rules applies 'cyclically', first operating on the most deeply embedded clause, then the next most deeply embedded, and so on, working 'up the tree' until they apply on the root clause, the entire generalized P-marker.
  - (28) Thus, singularly transformations apply to constituent clauses 'before' they are embedded, and to matrix clauses 'after' embedding has taken place. "The ordering possibilities that are permitted by the theory of Transformation-markers but apparently never put to use are now excluded in principle." (Chomsky 1965:135)
  - (29) Since minimalism returns to generalized transformations, in fact giving them even more prominence since all structure building is done by them instead of by PS rules, we must reconsider the Aspects argument against them.
  - (30) Recall that Chomsky argued that the Aspects model, lacking generalized transformations, excluded certain undesired interactions between generalized and singularly transformations (basically, anti-cyclic derivations).
  - (31) However, on closer inspection, it was not actually elimination of generalized transformations that had this limiting effect. Rather, it was the constraint that transformations operate bottom up, starting on the most deeply embedded clause and proceeding cyclically up the tree.
  - (32) Chomsky (1993) observes that a condition with the same effect can be imposed on the operation of generalized transformations (External Merge) and their interaction with singularly transformations (Internal Merge):
  - (33) Operations must extend their target (the Extension Condition), later subsumed under a general No Tampering Condition), and observes that this yields a version of the strict cycle. This guarantees the same sort of monotonic derivations as those permitted by Chomsky (1965).
  - (34) More recently Chomsky (2008) has proposed a No Tampering Condition for Merge whereby Merge cannot affect the internal structure of the two syntactic objects it combines. This has the same effect as the Extension Condition with respect to strict cyclicity.
  - (35) The one remaining Aspects argument against generalized transformations can also be straightforwardly addressed.
  - (36) Chomsky had argued that eliminating generalized transformations yields a simplified theory, with one class of complex operations jettisoned in favor of an expanded role for a component that was independently necessary, the phrase structure rule component.
  - (37) This was a very good argument. But since then, the transformational component has been dramatically restricted in its descriptive power. In place of the virtually unlimited number of highly specific transformations available under the theories of the 1950's and early 1960's, we have Merge. The complex apparent results come not from complex

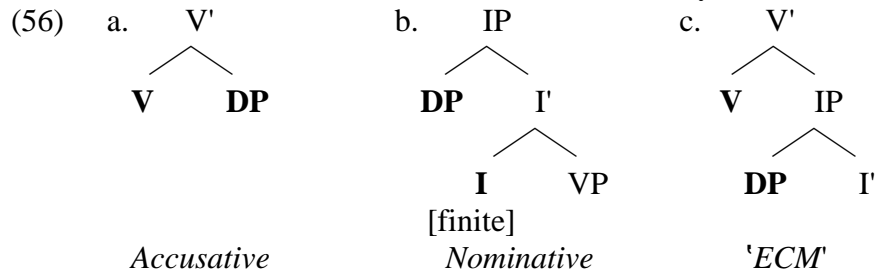
transformations, but from the interactions of this very simple one with very general constraints on the operation of transformations and on the ultimate derived outputs.

- (38) The 1965 argument can then be reversed on itself: Eliminate phrase structure rules, in the culmination of a program initiated in Stowell (1981).
- (39) Given the severely restricted inventory of elementary operations available, recent work in Minimalism suggests that 'single cycle' derivations would be the ideal for a computational architecture of human language.
- (40) Under the single cycle view, there is a single cyclic derivation, beginning with a selection of items from the lexicon, which builds the structure, successively transforms it, and periodically (at phases) sends information to the phonological and semantic interfaces.
- (41) In such a derivation syntactic, phonological, and semantic rules are interleaved in their application to linguistic structures. This stands in marked contrast to the classic principles and parameters 'Y-model':



- (43) In this model, a transformational cycle links an internal level D-structure to another internal level S-structure. A subsequent phonological cycle connects S-structure to the interface level PF.
- (44) Parallel to this latter cycle, a ('covert') transformational cycle relates S-structure to LF.
- (45) This multi-cycle model depends on the existence of internal levels of representation (i.e., beyond interface levels), which are prohibited under minimalist assumptions.
- (46) "This computational architecture, if sustainable, seems to be about as good as possible. S-structure and LF are no longer formulable as levels, hence disappear along with D-Structure, and computations are reduced to a single cycle" (Chomsky (2005, p.18))
- (47) The modern source of the single cycle concept is the Multiple Spell-Out proposal (Uriagereka (1996), Uriagereka (1999)) which Uriagereka (1999, p.276) suggests approximates the cyclic Spell-Out proposal in Chomsky (2000).
- (48) Under Multiple Spell-Out the operation of Spell-Out is itself cyclic, applying at several points during the derivation rather than just one.
- (49) This makes it impossible to separate a derivation into an overt vs. a covert part, as is possible in a derivation that has only one point where Spell-Out applies.
- (50) It also eliminates PF and LF as levels of representation in the technical sense, leaving no actual levels.
- (51) The computational architecture that results from cyclic Spell-Out is reminiscent of a proposal in Bresnan (1971), which both Uriagereka and Chomsky cite as an early (if only partial) antecedent.
- (52) Bresnan argues on empirical grounds that at least one class of phonological rules, those assigning stress patterns to sentences-and in particular, the Nuclear Stress Rule (NSR) of Chomsky and Halle (1968), must apply after the syntactic transformations in each cyclic domain. Therefore the NSR must be part of the syntactic cycle rather than part of a separate phonological cycle.

- (53) On the semantic side, Jackendoff (1969) and Jackendoff (1972) proposed that the rules assigning coreference relations to pronouns, reflexives, and null subjects of non-finite complements apply, just like Bresnan's version of the NSR, at the end of each syntactic cycle.
- (54) And Lasnik (1972) and Lasnik (1976) argued that scope of quantifiers and negation is assigned in that same fashion.
- (55) In the GB framework, *government* was a central structural relation. One of its major roles was in determining configurations of Case licensing or assignment. A minimalist desideratum is the elimination of such arbitrary relations.



- (57) Chomsky (1991) and Lasnik and Saito (1991) suggested that all Case licensing takes place in the core X-bar theoretic Spec-Head configuration, with (56)a and (56)c reduced to something like (56)b via movement to a Spec. This was a step toward minimizing the role of government.
- (58) More recently, Chomsky (2000), Chomsky (2001) proposed that feature licensing actually doesn't require movement at all.
- (59) Rather, a head (with a particular feature) probes for a goal (with a corresponding feature) in its domain.
- (60) If it finds one, and there is no intervener, Agree takes place.
- (61) The GB program very successfully approached human language 'from above', developing detailed proposals addressing the problem of explanatory adequacy (language acquisition).
- (62) Minimalism attempts to go beyond explanatory adequacy, approaching human language 'from below'.
- (63) There are many ways that the language faculty could, in principle, be organized so that language acquisition is easy. Why is it organized exactly the way it is?
- (64) Current minimalism seeks 'third factor' explanations, explanation that follow from general properties of computational systems and physical systems.
- (65) Merge and minimal search are plausible candidates for third factor properties of human language.
- (66) [If there is time, we will explore to what extent we can do without many GB constructs, while still maintaining reasonable descriptive adequacy – government, D-structure, S-structure, indices, ...]

## References

- Bresnan, Joan W. 1971. Sentence stress and syntactic transformations. *Language* 47: 257-281.
- Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms. Harvard University, Cambridge, Mass. and MIT, Cambridge, Mass., .[Revised 1956 version published in part by Plenum, New York, 1975; University of Chicago Press, Chicago, 1985].
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT press.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and parameters in comparative grammar*, ed. Robert Freidin, 417-454. Cambridge, Mass.: MIT Press. [Reprinted in Noam Chomsky, *The minimalist program*, 129-166. Cambridge, Mass.: MIT Press, 1995].
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge, Mass.: MIT Press. [Reprinted in Noam Chomsky, *The minimalist program*, 167-217. Cambridge, Mass.: MIT Press, 1995].
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory*, ed. Robert Freidin, Carlos Otero, and Maria-Luisa Zubizarreta. Cambridge, Mass.: MIT Press.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Fillmore, Charles J. 1963. The position of embedding transformations in a grammar. *Word* 19: 208-231.
- Jackendoff, Ray. 1969. *Some rules of semantic interpretation for English*. Doctoral dissertation, MIT, Cambridge, Mass.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
- Lasnik, Howard. 1972. *Analyses of negation in English*. Doctoral dissertation, MIT, Cambridge, Mass.
- Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2: 1-22. [Reprinted in *Essays on anaphora*, Howard Lasnik, 90-109. Dordrecht, Kluwer, 1989 ]
- Lasnik, Howard and Mamoru Saito. 1991. On the subject of infinitives. In *Papers from the 27th Regional Meeting of the Chicago Linguistic Society, Part I: The general session*, ed. Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, 324-343. Chicago Linguistic Society, University of Chicago, Chicago, Ill. [Reprinted in Howard Lasnik, *Minimalist analysis*, 7-24. Oxford: Blackwell, 1999].
- Stowell, Tim. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT, Cambridge, Mass.
- Uriagereka, Juan. 1996. Formal and substantive elegance in the Minimalist Program. In *The role of economy principles in linguistic theory*, ed. Chris Wilder, Hans-Martin Gärtner, and Manfred Bierwisch. Berlin: Akademie-Verlag.
- Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 251-282. Cambridge, Mass.: MIT Press.