

The First Half Century of Generative Grammar: Some Recurrent Themes

Howard Lasnik
University of Maryland
lasnik@umd.edu

I. The deterministic nature of derivations

Initially, and for quite some time, there were both optional and obligatory transformations, hence, branch points in transformational derivations. Thus, we find in Chomsky (1957) such transformations as the following, with the indicated specifications:

- (1)a Passive - optional
- b Number Transformation - obligatory
- c T_{not} - optional
- d T_q - optional
- e T_w - optional and conditional on T_q [The additional specification on this one is also very interesting. I will return to it.]

From the seem initial phrase marker, we get a multitude of sentences, depending on which optional transformations are selected:

- (2) John saw Mary
 Mary was seen by John
 Didn't John see Mary
 Was Mary seen by John
 etc.

With this sort of theory (many transformations, each explicitly marked optional or obligatory) serious learnability problems arise:

- (3) Suppose that there are N transformations in the grammar and suppose that the child has figured out all those N transformations. Imagine that now all the child has to figure out is for each rule whether it is optional or obligatory. In that case the child has to consider 2^N grammars. Each one of the N rules has two possibilities (optional and obligatory), and, by hypothesis, these possibilities are independent of those for any other rule, hence $2 \times 2 \times 2 \times \dots = 2^N$. As N gets big, as it indeed does in this theory, 2^N gets exponentially bigger.

Alongside this quantitative learnability problem, there is a qualitative problem as well:

- (4) How does the child know if a rule is obligatory? To determine that a rule is obligatory we, as linguists, use “negative data”, the information that a certain sentence is

unacceptable. But the child is not presented with negative data in any systematic fashion. Thus, if the child mistakenly guesses optional for an actually obligatory rule, recovery from the error is exceedingly problematic.

There are two salient solutions to this problem: All transformations are actually optional; or all are actually obligatory. The GB approach (explicitly so in Chomsky and Lasnik (1977)) was the former:

- (5) All Ts are optional, apparent obligatoriness following from surface filters.

However, almost from the earliest days of generative grammar, there were qualms about optional transformations:

- (6) "An obvious decision is to consider minimization of the optional part of the grammar to be the major factor in reducing complexity." Chomsky (1958/1962, p.154)

Chomsky (1965) extended this line of reasoning, suggesting that optional transformations might be eliminated:

- (7) "...it has been shown that many of the optional singular transformations of Chomsky (1955,1957,1962) must be reformulated as obligatory transformations, whose applicability to a string is determined by presence or absence of a certain marker in the string." p.132

- (8) "This was pointed out by Lees (1963) for the negation transformation, and by Klima (personal communication) for the question transformation ... In fact, it is also true of the passive transformation [on the account Chomsky presented in Aspects].

- (9) Following Katz and Postal (1964), Chomsky goes on to propose that "transformations cannot introduce meaning-bearing elements ..."

<The Inclusiveness Condition of Chomsky (1995, p.228) is an obvious extension of this idea.>

- (10) "These observations suggest that we attempt to formulate a more general condition ... that "nonstylistic transformations" are all signaled by optional markers drawn from a fixed, universal, language-independent set." p.223 fn.3

These concepts eventually emerged as the Principles and Parameters notion of 'Last Resort':

- (11) "... movement is a kind of 'last resort'. An NP is moved only when this is required, either because it is a wh-phrase that must appear in operator position (at least an [sic] LF), or in order to escape a violation of some principle: the Case Filter, ... or a principle ... requiring nongovernment of the empty category PRO." p.143

Chomsky and Lasnik (1993) adopt the same notion, incorporated into 'economy of derivation', which became a cornerstone of Minimalism:

- (12) "The principle of economy of derivation requires that computational operations must be driven by some condition on representations, as a 'last resort' to overcome a failure to meet such a condition." p.28

Chomsky (1993) formulates the Last Resort principle as 'Greed':

- (13) Movement of α to β must be for the satisfaction of formal requirements of α .

True optionality is thus minimized, and derivations become much more deterministic. Another sort of problem emerges, though. Chomsky's transformational theories have always tended to have a Markovian character - the next step in the derivation depends solely on the current phrase marker, yet for movement to obey Greed, knowledge of the ultimate derived structure is required, a situation sometimes called 'Look-ahead'. This is most strikingly true in the case of successive cyclic movement:

- (14) Mary is believed [t' to be likely [t' to be elected t]]

It is only the final step of movement that satisfies the Case requirement of *Mary*. The prior steps are permitted because they make the last step possible.

The mirror image of this non-Markovian Look-ahead occurred in Chomsky's early work, where 'Look-back' was occasionally used, as in (1)e above, repeated here:

- (15) Tw - optional and conditional on T_q

T_w is the ancestor of wh-movement (though it also created wh-phrases out of plain NPs before it moved them). T_q is Subject-Aux Inversion (Movement of finite Infl to Comp in more recent theories). The non-Markovian statement on the rule is to guarantee that T_w applies only if SAI had already applied:

- (16)a Who will the students see
b *Who the students will see

Interestingly, this interaction remains problematic half a century later.

Returning, now, to the Look-ahead problem of Greed-based movement, Chomsky (1995) addresses this by turning Greed upside down:

- (17) Movement of α to β must be for the satisfaction of formal requirements of β .

Metaphorically, instead of α moving to β , β 'attracts' α . But what are the needs of the successive subject positions in successive cyclic NP movement?

- (18) The Extended Projection Principle (EPP) (here called by Chomsky 'principle P') "is the structural requirement that certain configurations ... must have subjects..." Chomsky (1981, p.27)

- (19)a. It seems that John is here
 b. *Seems that John is here
- (20) This does not entirely follow from θ -theory, since even when the predicate has no subject θ -role to assign, a subject must nonetheless be present, at least in one class of languages. "...the subject of a clause is obligatory in English and similar languages." [p.40]
- (21) Chomsky (1982) introduces the name 'Extended Projection Principle', since the requirement goes beyond anything demanded by the Projection Principle, "which states informally that the θ -marking properties of each lexical item must be represented categorially at each syntactic level...". [p.8]
- (22) Given that finite Infl is a Case 'assigner' (nominative Case), Fukui and Speas (1986) propose that the effects of the EPP actually follow from a more general requirement that a Case assigner must assign its Case. (19)b is out because Infl is unable to assign its Case.
- (23) But there are situations where neither θ -theory nor Case theory demand a subject, yet one is apparently still required (even if the result is ungrammatical; i.e., with or without a (pleonastic) subject, the sentences are bad).
- (24) *the belief [to seem [Peter is ill]]
 (25) *[To seem [Peter is ill]] is widely believed
 (26) *John has conjectured [to seem [Peter is ill]] Boskovic (1997)

The roots of the EPP can be traced even further back. The basic phrase structure rules always had optional complements for V, but an obligatory NP dominated by S:

- (27)a $S \rightarrow NP VP$
 b $VP \rightarrow V (\dots)$

This asymmetry carried over to the important notion 'subcategorization' introduced in Chomsky (1965):

- (28) "Corresponding to each string dominated by VP, there is a strict subcategorization of Verbs. On the other hand, Verbs are not strictly subcategorized in terms of types of Subject NPs ..." p.96

That is, some verbs take objects, some don't; some are ditransitive, some aren't; but all verbs take subjects (even, as noted above, when there is no semantic reason that there should be one). There have been numerous attempts to eliminate the EPP, or to deduce its effects from more natural principles (see, for example, Epstein and Seely (2006) for extensive discussion), but a residue of the EPP still remains.

II. Constraints on movement

In early generative grammar, transformations were far more specific than in later developments of the theory. These later developments usually had the effect of factoring properties out of specific transformations. One rather striking example is locality, of which there is virtually no discussion in Chomsky (1955). The only hint is in the p. 437 discussion of a certain restriction on the WH transformation, as in the following unacceptable example:

(29) Whom did your interest in seem to me rather strange

Surprisingly from a modern perspective, Chomsky suggested incorporating the constraint into the transformation itself. It was only in Chomsky (1964a) that we find the idea of formulating general constraints on the operation of transformations, and, of course, only in Ross (1967) that we find a full blown development of the idea.

In Chomsky (1964a), we find a relatively modern looking wh-movement transformation, which is both simple and quite general:

(30) $Y - Wh+X - Z \Rightarrow Wh+X - Y - Z$

As Chomsky notes, this generality raises potential problems of overgeneration. Some of these problems are addressed by constraints on movement. It is interesting to note that at least some of the discussions of movement constraints occasionally hint at a connection to the deterministic nature of derivations. That is, it is suggested that the constraints resolve potential ambiguity of application.

(31) "... although several Noun Phrases in a sentence may have *Wh* attached to them, the operation [(30)] must be limited to a single application to each underlying terminal string. Thus we can have "who saw what?", "you met the man who saw what?", "you read the book that who saw?", "you saw the book which was next to what?", etc., but not "who what saw?", "you saw the book which which was next to" (as a declarative), and so on, as could arise from multiple applications of this rule. These examples show that [(30)] cannot apply twice to a given string ... " p.43

Chomsky provides several additional arguments that this constraint is necessary. And again, in a foreshadowing of modern concerns, Chomsky raises questions about the nature of the constraint:

(32) "The constraint that [(30)] may not reapply to a given string is thus necessary if the grammar is to achieve descriptive adequacy. Once again, to achieve the level of explanatory adequacy, we must find a principled basis, a general condition on the structure of any grammar, that will require that in the case of English the rule [(30)] must be so constrained. Various suggestions come to mind, but I am unable to formulate a general condition that seems to me entirely satisfying." p.45

In later developments, Subjacency, Superiority, and Relativized Minimality were proposed as more general constraints from which these specific cases follow.

The A-over-A constraint (but not by that name) is also first suggested in Chomsky (1964b), along with the proviso that the constraint is too strong (as discussed in great detail later by Ross (1967)).

- (33)a What would it be difficult for me to understand
b *What would for me to understand be difficult

(34) "... in the illegitimate case the Noun Phrase to be preposed is contained within a Noun Phrase [This assumes that subjects are always NPs.], while in the legitimate case, it is not. However, the condition that a Noun Phrase contained within a Noun Phrase is not subject to[(30)], though quite plausible and suggested by many examples, is apparently somewhat too strong, as we can see from such, to be sure, rather marginal examples as 'who would you approve of my seeing?', 'what are you uncertain about giving to John?', 'what would you be surprised by his reading?', etc." p.46

Ross (1967) argued against both of the constraints Chomsky suggested, proposing, in their place, a set of 'island' constraints on movement (or, more specifically, on the 'chopping' part of movement, as Ross decomposed movement into copying and chopping).

Some of Ross's constraints:

Complex NP Constraint (modified version of a constraint attributed to Ed Klima)

- (35) No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.
(36)a *The man who I read a statement which was about is sick.
b The man who I read a statement about is sick.

Coordinate Structure Constraint

- (37) In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.
(38)a *What sofa will he put the chair between some table and?
b *Whose tax did the nurse polish her trombone and the plumber compute?

Left Branch Condition

- (39) No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule.
(40)a The boy whose guardian's employer we elected president ratted on us.
b *The boy whose guardian's we elected employer president ratted on us.
c *The boy whose we elected guardian's employer president ratted on us.

Sentential Subject Constraint

(41) No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S.

<This again assumes that subjects are always NPs, so that sentential subjects are dominated by NP.>

- (42)a The teacher who the reporters expected that the principal would fire is a crusty old fizzlebotch.
b *The teacher who that the principal would fire was expected by the reporters is a crusty old fizzlebotch.
c The teacher who it was expected by the reporters that the principal would fire is a crusty old fizzlebotch.

In Chomsky (1973), Chomsky for the first time explored constraints on the operation of transformations in detail. He maintained versions of the the two constraints of Chomsky (1964a) and, probably most significantly, proposed Subjacency as a constraint intended to unify some of Ross's island constraints.

(43) "... if X is superior to Y in a phrase marker P [roughly, if X asymmetrically c-commands Y], then Y is '**subjacent**' to X if there is at most one cyclic category $C \neq Y$ such that C contains Y and C does not contain X. Thus, if Y is subjacent to X, either X and Y are contained in all the same cyclic categories or they are in adjacent cycles." p.247

(44) No [movement] rule can involve X, Y, X superior to Y if Y is not subjacent to X.

This had the major new consequence that apparent unbounded movement was actually constituted of a series of short movements. Just like NP-movement, wh-movement must be successive cyclic.

About the A-over-A constraint, one of the constraints Chomsky maintains, there are several comments indicative of the constraint reducing derivational alternatives:

A-over-A

(45) If a transformation applies to a structure of the form

$[\alpha [A \dots] \dots]$

where α is a cyclic node, then it must be so interpreted as to apply to the maximal phrase of the type A.

(46) "Notice that the condition [(45)] does not establish an absolute prohibition against transformations that extract a phrase of type A from a more inclusive phrase of type A. Rather, it states that if a transformational rule is nonspecific with respect to the configuration defined, it will be interpreted in such a way as to satisfy the condition. Conditions on Ts." p.235

As for (31), Chomsky (1973) decomposed it, part of it (the ban on extraction from an embedded question) falling under Subjacency. For the residue, Chomsky proposed the Superiority condition:

- (47) No rule can involve X, Y in the structure
 ...X... [_α... Z...-WYZ ...] ...
 where the rule applies ambiguously to Z and Y and Z is superior to Y.
- (48) "The condition requires that a rule must select the superior term where that rule is ambiguous in application, that is, where the structure given in [(47)] will satisfy the structural condition defining the rule in question with either Z or Y selected as the factor satisfying a given term of this condition. Like the A-over-A Condition, [(47)] restricts the ambiguity of rule application." p.246

The proper formulation of Subjacency became a major research question (and continues to be one). In part because of arguments in Williams (1974), the notion 'cyclic node' became much less clear. Subsequent versions of Subjacency thus generally referred to 'bounding nodes' or 'barriers', the bounding nodes being S (IP) and NP. Chomsky (1986) was a bold attempt at a new principled theory of bounding nodes. The core idea is that ALL XPs are potentially barriers, but that an XP that is the complement of a lexical head (V, N, A, maybe P) is not a barrier. This gives the subject-object asymmetry discussed by Chomsky (1973), since object is complement of V:

- (49)a Who did [you read [stories about t]]
 b *Who did you go home [because Mary mentioned t]

This new 'Barriers' theory also accounts for the observation of Huang (1982) that extraction out of 'adjuncts' (including adverbial modifiers) is barred:

- (50) *Who did you go home [because Mary mentioned t]

Sadly, to handle the full range of relevant facts, the theory became extremely complicated.

Chomsky's next detailed proposal came a decade and a half later, in Chomsky (2000) and Chomsky (2001). 'Barrier' is replaced by 'Phase', where the phases are vP and CP.

- (51) "... the phases are 'propositional': verbal phrases with full argument structure and CP with force indicators ..."
 Chomsky (2001, p.12)

Subjacency is then replaced by the Phase-Impenetrability Condition:

- (52) For phase HP with head H,
 The domain of H is not accessible to operations outside HP; only H and its *edge* are accessible to such operations,
 the edge being the residue outside H', either specifiers or elements adjoined to HP.

This new approach to barrierhood meshes with a new approach to derivation: The syntactic structure is built strictly bottom up (à la the generalized transformations of Chomsky (1955)), with movement processes interspersed with structure building ones. Further, phonological and semantic interpretation is performed cyclically (as in Uriagereka (1999)), phase by phase. Once a derivational phase is reached, material from the preceding phase is "handed over" to the interface components. Islandhood is then an inevitable consequence of this multiple spell-out. Once a structure is sent for phonological interpretation, it is frozen.

I will end by pointing out that one island phenomenon explored early on (by Ross (1969)) and much more recently (by Merchant (2001)) raises a profound problem. Ross observed that island violations are dramatically ameliorated by deletion (S deletion 'Sluicing'):

- (53) I believe that he bit someone, but they don't know who (I believe that he bit)
- (54)a *I believe the claim that he bit someone, but they don't know who I believe the claim that he bit [Complex NP Constraint, noun complement]
 b(??)I believe the claim that he bit someone, but they don't know who
- (55)a *Irv and someone were dancing together, but I don't know who Irv and were dancing together [Coordinate Structure Constraint]
 b(??)Irv and someone were dancing together, but I don't know who
- (56)a *She kissed a man who bit one of my friends, but Tom doesn't realize which one of my friends she kissed a man who bit [Complex NP Constraint, relative clause]
 b(??)She kissed a man who bit one of my friends, but Tom doesn't realize which one of my friends
- (57)a *That he'll hire someone is possible, but I won't divulge who that he'll hire is possible [Sentential Subject Constraint]
 b (??)That he'll hire someone is possible, but I won't divulge who

As Ross already pointed out, this phenomenon seems to demand 'globality', in violation of the assumed Markovian character of derivations. Further, and perhaps even more problematically, if islandhood results from material being frozen in place, how is island repair possible at all? This question is a hot current topic, being explored by, among many others, Merchant (2001), Merchant (In press), Lasnik (2001), Fox and Lasnik (2003), and Fox and Pesetsky (2003).

References

- Boskovic, Zeljko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.
- 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. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1958/1962. A transformational approach to syntax. In *Proceedings of the Third Texas conference on problems of linguistic analysis in English*, ed. Archibald A. Hill, 124-158. University of Texas Press, University of Texas, Austin. [Reprinted in *The structure of language: Readings in the philosophy of language*, eds Jerry A. Fodor and Jerrold J. Katz, 211-245. Englewood Cliffs, N. J.: Prentice-Hall, 1964].
- Chomsky, Noam. 1964a. *Current issues in linguistic theory*. The Hague: Mouton.
- Chomsky, Noam. 1964b. The logical basis of linguistic theory. In *Proceedings of the Ninth International Congress of Linguistics*, ed. Horace Lunt, 914-978. The Hague: Mouton. Reprinted as "Current Issues in Linguistic Theory." In *Structure of Language*, edited by Fodor and Katz. New York: Prentice-Hall, 1964.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT press.
- Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.

- 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. Categories and transformations. In *The minimalist program*, 219-394. 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 and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 11: 425-504. [Reprinted in Howard Lasnik, *Essays on restrictiveness and learnability*, 42-124. Dordrecht: Kluwer, 1990]
- Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, Vol. 1, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506-569. Berlin: Walter de Gruyter. [Reprinted in Noam Chomsky, *The minimalist program*, 13-127. Cambridge, Mass.: MIT Press, 1995].
- Epstein, Samuel D. and T. Daniel Seely. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.
- Fox, Danny and Howard Lasnik. 2003. Successive cyclic movement and island repair: The difference between Sluicing and VP Ellipsis. *Linguistic Inquiry* 34: 143-154.
- Fox, Danny and David Pesetsky. 2003. Cyclic linearization and the typology of movement. Ms. MIT, Cambridge, Mass.
- Fukui, Naoki and Margaret Speas. 1986. Specifiers and projection. In *MIT working papers in linguistics* 8, 128-172.
- Huang, C. -T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT, Cambridge, Mass.
- Katz, Jerrold J. and Paul Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT Press.
- Lasnik, Howard. 2001. Derivation and representation in modern transformational syntax. In *Handbook of syntactic theory*, ed. Mark Baltin and Chris Collins, 62-88. Oxford: Blackwell.
- Lees, Robert B. 1963. *The grammar of English nominalizations*. The Hague: Mouton. (Originally published 1960 by Indiana University. Research Center in Anthropology, Folklore, and Linguistics.)
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason. In press. Variable island repair under ellipsis. In *Topics in ellipsis*, ed. Kyle Johnson. Cambridge: Cambridge University Press.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT, Cambridge, Mass. Published as *Infinite syntax!* Norwood, N.J.: Ablex (1986).
- Ross, John Robert. 1969. Guess who? In *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, ed. Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252-286. Chicago Linguistic Society, University of Chicago, Chicago, Ill.
- Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 251-282. Cambridge, Mass.: MIT Press.
- Williams, Edwin. 1974. *Rule ordering in syntax*. Doctoral dissertation, MIT, Cambridge, Mass.