Derivation vs. Representation in Modern Transformational Grammar

Howard Lasnik
University of Connecticut

I. Locality of movement: Subjacency

(1) Who do you think Mary said John likes
(2) ?*Who did you ask whether Mary knows why John likes

(3) Chomsky (1973), rejecting earlier views (including the highly influential Ross (1967)) proposed that long distance movement is never possible.
(4) Movement across more than one bounding node is prohibited by Subjacency and S is a bounding node.

(5) [$_{S'}$ Who [$_{S}$ you think [$_{S'}$ t [$_{S}$ Mary said [$_{S'}$ t [$_{S}$ John likes t]]]]]]

(6) Movement is via 'Comp', given the phrase structure:
(7) S' → Comp S

(8) [$_{S'}$ Who [$_{S}$ you ask [$_{S'}$ whether [$_{S}$ Mary knows [$_{S'}$ why [$_{S}$ John likes t]]]]]]
(9) Who do you think that Mary said that John likes

(10) (9) has the grammaticality status of (1), but does not seem relevantly different from (2) in its structure.

(11) [$_{S'}$ Who [$_{S}$ you think [$_{S'}$ that [$_{S}$ Mary said [$_{S'}$ that [$_{S}$ John likes t]]]]]]
(12) In the course of the syntactic derivation, the structure of (9) is identical to that of (1), namely (5). Thus, movement proceeds via successive Comps in both derivations.
(13) But in (9), at a very late level of derivation the non-matrix Comps are spelled out as that.
(14) At that late level, Subjacency appears to be violated, even though in the course of the derivation every step is legitimate.
(15) This is the paradigmatic type of situation Chomsky frequently alludes to in his recent writings when he argues for a derivational approach to syntax.

II. Locality of movement: ECP

(16) ?Which problem do you wonder whether Mary solved
(17) *How do you wonder whether Mary solved the problem

(18) Huang (1982), extending ideas of Chomsky (1981), appealed to the Empty Category Principle (ECP), a constraint independent of Subjacency demanding locality between a (non-argument) trace and its antecedent.
(19) (17) is then worse than (16) because (16) violates only Subjacency while (17) violates both Subjacency and the ECP.

(20) How do you think that Mary solved the problem

(21) *How do you wonder whether John said (that) Mary solved the problem
(22) ??Which problem do you wonder whether John said (that) Mary solved

(23) Even intermediate traces must be 'properly governed'.
(24) How [$_{S}$ do you wonder [$_{S'}$ whether [$_{S}$ John said [$_{S'}$ t' [$_{S}$ Mary solved the problem t]]]]]
(25) The initial trace t is properly governed by the intermediate trace t'. But the intermediate trace is too distant from its antecedent How, causing a violation of the ECP.
(26) How [$_{S}$ do you wonder [$_{S'}$ whether [$_{S}$ John said [$_{S'}$ that [$_{S}$ Mary solved the problem t]]]]]
(27) How did you say that Mary solved the problem

(28) The initial trace must be established as properly governed at a point in the derivation where the that is not present.
(29) That point might be before that is inserted.
(30) Alternatively, the point might be after that is inserted and subsequently deleted in the LF component.
(31) Under either alternative, the question is how we can resolve the apparent contradiction between the unacceptable (21) and the acceptable (27).
(32) How do you think [$_{S'}$ [$_{S}$ Mary solved the problem t]]
(33) Chomsky (1986) rejects the 'representational' stipulation that gamma marking is only at levels.
(34) [$_{CP}$ [$_{C'}$ do [$_{IP}$ you wonder [$_{CP}$ whether [$_{IP}$ John said [$_{CP}$ how [$_{C'}$ (that) [$_{IP}$ Mary solved the problem t]]]]]]

(35) Adjuncts must be fully represented.

(36) [$_{CP}$ How [$_{C'}$ do [$_{IP}$ you wonder [$_{CP}$ whether [$_{IP}$ John said [$_{CP}$ t' [$_{C'}$ (that) [$_{IP}$ Mary solved the problem t]]]]]]

(37) ??Which problem do you wonder whether John said (that) Mary solved
(38) [$_{CP}$ Which problem [$_{C'}$ do [$_{IP}$ you wonder [$_{CP}$ whether [$_{IP}$ John said [$_{CP}$ t' [$_{C'}$ (that) [$_{IP}$ Mary solved t]]]]]]]

(39) "An expression...is a Subjacency violation if its derivation forms a starred trace. It is an Empty Category Principle (ECP) violation if, furthermore, this starred trace remains at LF..." Chomsky and Lasnik (1993)
(40) Significantly, even on this 'derivational' analysis, the ECP is actually not entirely derivational.

(41) Every element in an LF representation must be 'legitimate'.
(42) a. Arguments: each element is in an A-position.
b. Adjuncts: each element is in an $\bar{A}$-position.
c. Lexical elements: each element is in an X^0 position.
d. Predicates, possibly predicate chains if there is predicate raising.
d. Operator-variable constructions, each a chain (α_1,α_2), where the operator α_1 is in an $\bar{A}$-position and the variable α_2 is in an A-position.

Chomsky (1991)

(43) As before, the approach is derivational, in that the marking of a trace as being in violation of the ECP crucially takes place in the course of the derivation. Yet it is representational in the way that a derivational violation can be remedied. If no offending trace remains at the level of LF, the resulting structure is not an ECP violation, even if on-line it was.

(44) Subjacency remains strictly derivational: a long movement, even of an argument, causes some degradation of the sentence.

(45) Mary hired someone.
Tell me who ~~Mary hired~~.

(46) I heard that Mary hired someone.
Tell me who ~~you heard that Mary hired~~.

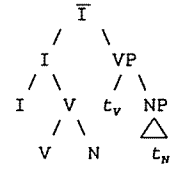
(47) I resigned because Mary hired someone.
?*Tell me who you resigned because Mary hired.
?Tell me who ~~you resigned because Mary hired~~.

(48) For both ECP and Subjacency, an on-line violation can be improved by a later operation that results in a change in the ultimate representation, LF in the first case, PF in the second.

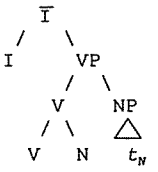
(49) Chomsky argues for the reality of syntactic derivations by appealing to virtually the opposite state of affairs: situations where the ECP is <u>satisfied</u> on-line, but where the ultimate LF representation appears to be in violation, yet the resulting sentence is <u>good</u>.

(50) "Viewed derivationally, computation typically involves simple steps expressible in terms of natural relations and properties, with the context that makes them natural 'wiped out' by later operations, hence not visible in the representations to which the derivation converges. Thus, in syntax, crucial relations are typically local, but a sequence of operations may yield a representation in which the locality is obscured. Head movement, for example, is narrowly 'local,' but several such operations may leave a head separated from its trace by an intervening head. This happens, for example, when N incorporates to V, leaving the trace t_N and the $[_V$ V-N] complex then raises to I, leaving the trace t_V: the chain (N, t_N) at the output level violates the locality property, and further operations (say, XP-fronting) may obscure it even more radically, but locality is observed in each individual step." [Chomsky (1995,pp. 223-224)]

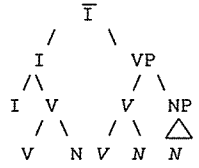
(51)



(52)



(53)



(54) If N 'antecedent governs' its trace in (52), presumably the intermediate *N* likewise antecedent governs the initial trace *N* in (53).

(55) "It is generally possible to formulate the desired result in terms of outputs. In the head movement case, for example, one can appeal to the (plausible) assumption that the trace is a copy, so the intermediate V-trace includes within it a record of the local N → V raising. But surely this is the wrong move. The relevant chains at LF are (N, t_N) and (V, t_V), and in these the locality relation satisfied by successive raising has been lost." Chomsky (1995, p.224)

III. Intermediate traces?

(56) Do chains contain intermediate traces at LF?
(57) Adjunct-argument island asymmetries, as discussed earlier.
(58) Reconstruction effects:
(59) Mary wondered which picture of himself Bill saw *t*
(60) Which picture of himself does Mary think that John said that Susan likes See Barss (1986).
(61) *Mary thinks that John said that Susan likes pictures of himself
(62) Mary thinks that John said that pictures of himself, Susan likes

(63) Are there really two arguments for intermediate traces?

(64) The argument based on the adjunct-argument asymmetry implicates an analysis under which intermediate traces of arguments are eliminated prior to the LF level (while intermediate traces of adjuncts remain). But the reconstruction effects implicating intermediate traces involved WH-argument movement.

(65) Chomsky's recent approach to reconstruction centers around traces (copies) visible at the LF level, a strongly representational approach.

(66) A more derivational alternative: Belletti and Rizzi (1988), where Condition A can be satisfied on-line.

(67) If we take the Binding Theory to consist not of conditions on form, but rather of interpretive principles, then this amounts to saying that there is no specific <u>level</u> of LF.

(68) Compare the Jackendoff (1972) theory of anaphora, which included interpretive rules operating at the end of each syntactic cycle.

(69) Similarly, Lasnik (1972) suggested cyclic interpretation of the scope of negation, and Lasnik (1976) extended this to the scope of other operators.

(70) On the PF side, Bresnan (1971) argued that the rule responsible for the assignment of sentence stress in English applies not at Surface Structure, as had been assumed, but at the end of each syntactic cycle.

(71) Epstein (1995) (see also Uriagereka (In press)) suggests that all interpretive information is provided on-line, in the course of the syntactic derivation. There is no level of LF per se.

IV. 'Strong feature' violations: PF, LF, or on-line?

(72) When movement is overt, it must have been forced to operate 'early' by some special requirement: a 'strong feature' that needs to be checked. Chomsky (1993;1994;1995)

(73)**A** A strong feature that is not checked in overt syntax causes a derivation to crash at PF. Chomsky (1993)

B A strong feature that is not checked (and eliminated) in overt syntax causes a derivation to crash at LF. Chomsky (1994)

C A strong feature must be eliminated (almost) immediately upon its introduction into the phrase marker. Chomsky (1995, ch.4)

(74) The PF theory, (73)**A**:
"...'strong' features are visible at PF and 'weak' features invisible at PF. These features are not legitimate objects at PF; they are not proper components of phonetic matrices. Therefore, if a strong feature remains after Spell-Out, the derivation crashes... Alternatively, weak features are deleted in the PF component so that PF rules can apply to the phonological matrix that remains; strong features are not deleted so that PF rules do not apply, causing the derivation to crash at PF."

(75) *John read what?

(76) The LF theory, (73)**B**:

(77)a "...Spell-Out can apply anywhere, the derivation crashing if a 'wrong choice' is made...If the phonological component adds a lexical item at the root, it will introduce semantic features, and the derivation will crash at PF. If the covert component does the same, it will introduce phonological features, and the derivation will therefore crash at LF...

b Suppose that root C (complementizer) has a strong feature that requires overt wh-movement. We now want to say that unless this feature is checked before Spell-Out it will cause the derivation to crash at LF to avoid the possibility of accessing C after Spell-Out in the covert component."

(78) Spell-Out: C [strong Q] John read what *LF

(79) Spell-Out: John read what
LF: C [strong Q] John read what *LF

(80) The on-line theory, (73)**C**:
"We ... define a strong feature as one that a derivation 'cannot tolerate': a derivation D→Σ is canceled if Σ contains a strong feature..."

(81) Ellipsis provides potential evidence for (**A**), if it is, as suggested by Chomsky and Lasnik (1993), a PF deletion process.

(82) Two instances: first Pseudogapping then Sluicing.

(83)a If you don't believe me, you will ø the weatherman

b I rolled up a newspaper, and Lynn did ø a magazine

c Kathy likes astronomy, but she doesn't ø meteorology
Levin (1978)

(84)a The DA proved Jones guilty and the Assistant DA will ~~prove~~ Smith ~~guilty~~

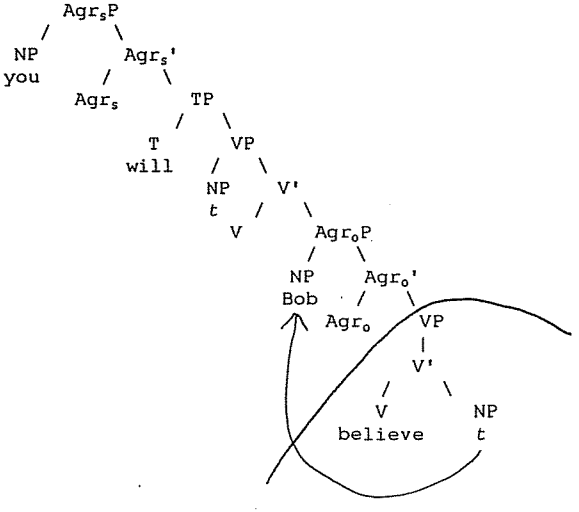
b ?John gave Bill a lot of money, and Mary will ~~give~~ Susan ~~a lot of money~~

(85) You might not believe me but you will Bob

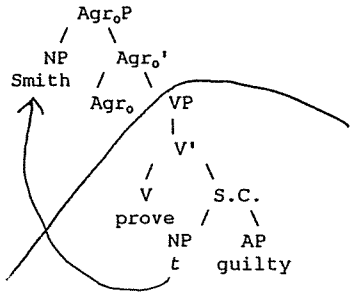
(86) NP-raising to Spec of Agr_o ('Object Shift') is overt in English. [Koizumi (1993), Koizumi (1995), developing ideas of Johnson (1991)]

(87) Pseudogapping as overt raising to Spec of Agr_o followed by deletion of VP. [Lasnik (1995b)]

(88)



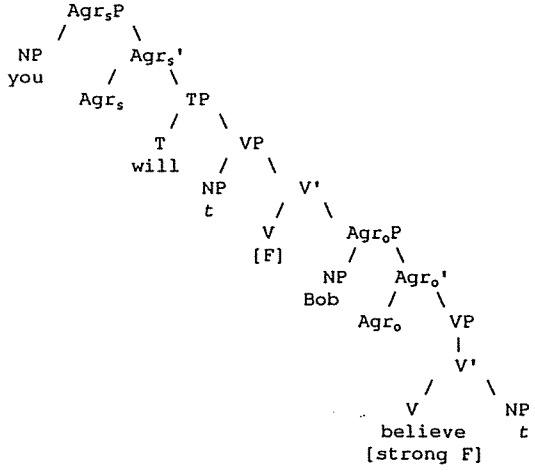
(89)



(90) *You will Bob believe

(91) *The Assistant DA will Smith prove guilty

(92)



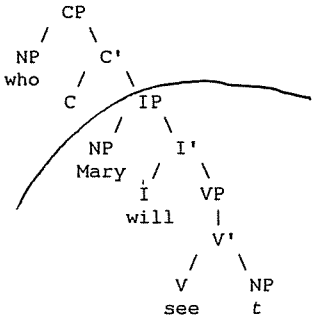
(93) Suppose the strong feature driving V-raising resides in the lexical V rather than in the higher 'shell' V. The strong feature of the verb must **either** be checked by overt raising to the shell V **or** be contained in an ellipsis site. PF deletion could eliminate the unchecked strong feature.

(94) Sluicing - WH-Movement followed by deletion of IP (abstracting away from 'split Infl' details). [Saito and Murasugi (1993), Lobeck (1990)]

(95) Speaker A: Mary will see someone.
Speaker B: I wonder who ~~Mary will see~~.

(96) Speaker A: Mary will see someone.
Speaker B: Who ~~Mary will see~~?

(97)



(98) *Who Mary will see?

(99) Who will Mary see?

(100) Suppose that in a matrix interrogative, it is Infl that has a strong feature, rather than C. The strong feature of Infl must **either** be checked by overt raising to the interrogative C **or** be contained in an ellipsis site. PF deletion could eliminate the unchecked strong feature.

V. Digression: On the overtness of English object-shift

(101) Infl-raising to C is uncontroversially overt in normal matrix interrogatives. NP-raising to Spec of Agr$_o$, on the other hand, is standardly assumed to be covert in English. Lasnik (1995a), based on Lasnik and Saito (1991) [see also Postal (1974) and Wyngaerd (1989)] and den Dikken (1995), argues that such movement is overt.

(102)a There is a man here
b There are men here

(103)a Many linguistics students aren't here
b There aren't many linguistics students here

(104)a Some linguists seem to each other [t to have been given good job offers]
b *There seem to each other [t to have been some linguists given good job offers]

(105)a No good linguistic theories seem to any philosophers [t to have been formulated]
b *There seem to any philosophers [t to have been no good linguistic theories formulated]

(106)a Some defendant$_i$ seems to his$_i$ lawyer [t to have been at the scene]
b *There seems to his$_i$ lawyer [t to have been some defendant$_i$ at the scene]

(107) "The operation Move...seeks to raise just F." Chomsky (1995)

(108) When movement is covert, hence only of formal features, the referential and quantificational properties needed to create new binding and scope configurations are left behind, so no such new configurations are created. Lasnik (1995a) (contra Chomsky (1995), at least in part)

(109) The DA questioned two men during each other's trials
(110)a The DA proved [two men to have been at the scene] during each other's trials
b *The DA proved [there to have been two men at the scene] during each other's trials

(111) The DA questioned noone during any of the trials
(112)a The DA proved [noone to be at the scene] during any of the trials
b *The DA proved [there to be noone at the scene] during any of the trials

(113) The DA questioned no suspect$_i$ during his$_i$ trial
(114)a The DA proved [no suspect$_i$ to be at the scene of the crime] during his$_i$ trial
b *The DA proved [there to be no suspect$_i$ at the scene of the crime] during his$_i$ trial

(115) One further argument: Given the feature movement theory of covert movement, if an instance of movement creates a new ellipsis configuration, that movement must be overt. (This is true whether ellipsis is PF deletion or LF copying.)

VI. Conceptual issues surrounding the strong feature controversy

(116) Possible arguments against the PF approach to strong features (73)**A**:
(117)a 'Look-ahead' is needed. At a given point in the overt portion of a derivation, it is necessary to inspect the PF representation to see whether Procrastinate can be evaded. [The LF approach (73)**B** shares this problem (if it is a problem).]
b The derivation of **John read what* in ? above, with covert insertion of C with a strong feature, won't be blocked.

(118) (73)**C** above, repeated here, is designed to eliminate the Look-ahead problem.
(119) A strong feature must be eliminated (almost) immediately upon its introduction into the phrase marker. Chomsky (1995, ch.4)
(120) "We...virtually derive the conclusion that a strong feature triggers an *overt* operation to eliminate it by checking. This conclusion follows with a single exception: covert merger (at the root) of a lexical item that has a strong feature but no phonological features..."

(121) *John read what

(122) To prevent this, covert insertion of strong features must be barred.
(123) α enters the numeration only if it has an effect on output.
(124) "Under [(123)], the reference set [for economy comparisons] is still determined by the numeration, but output conditions enter into determination of the numeration itself..."

(125) Look-ahead?

(126) "With regard to the PF level, *effect* can be defined in terms of literal identity...α is selected only if it changes the phonetic form.
(127) At the LF level the condition is perhaps slightly weaker, allowing a narrow and readily computable form of logical equivalence to be interpreted as identity."

(128) Clearly, covert insertion of a C will have no phonetic effect. Will it have an effect at the LF output?

(129) If it will, then covert insertion is allowed, and we generate (121) with structure (130):
(130) C [$_{IP}$ John read what]

(131) If it will not, then we generate (121) with structure (132):
(132) [$_{IP}$ John read what]
(133) (132) violates no morphological requirements, and, if C has no effect on output, then it should mean exactly *What did John read?*
(134) "...the interface representations (π,λ) are virtually identical whether the operation [covert insertion of strong features] takes place or not. The PF representations are in fact identical, and the LF ones differ only trivially in form, and not at all in interpretation."

(135) (73)**C** entails that strength is always a property of an 'attracting' head, never a property of the item that moves. The above analyses of Pseudogapping and Sluicing are incompatible with that proposal.

(136) There is a possible alternative analysis, based on the Chomsky (1995, ch.4) theory of pied-piping, particularly as explicated by Ochi (1997).
(137) "For the most part - perhaps completely - it is properties of the phonological component that require pied-piping. Isolated features and other scattered parts of words may not be subject to its rules, in which case the derivation is canceled; or the derivation might proceed to PF with elements that are 'unpronounceable,' violating FI."
(138) "Just how broadly considerations of PF convergence might extend is unclear, pending better understanding of morphology and the internal structure of phrases. Note that such considerations could permit raising without pied-piping even overtly, depending on morphological structure..."

(139) Matrix interrogative C might then contain the strong feature, with the matching feature of Infl raising overtly to check it. This leaves behind a phonologically defective Infl, which will cause a PF crash unless either pied-piping or deletion of a category containing that Infl (Sluicing) takes place.
(140) Similarly for the feature driving overt V-raising: it could be a strong feature of the **higher** V. Once the matching feature of the lower lexical V is 'attracted', the lower V becomes defective. A PF crash will be avoided if either pied-piping or deletion of a category containing the lower V (VP Deletion = Pseudogapping in the relevant instances) takes place.

(141) However, there is independent evidence **for** strong features residing in moving categories, and hence **against** the strictly derivational (73)**C**.

(142) For example, Boskovic (1997) shows that in Serbo-Croatian, WH-phrases have a strong focus feature: they all have to move overtly.

(143)a Ko šta gdje kupuje?
who what where buys "Who buys what where?"
*Ko kupuje šta gdje?
*Ko šta kupuje gdje?
*Ko gdje kupuje šta?

(144) Is syntax derivational or representational?

(145) Yes.

References

Barss, Andrew. 1986. *Chains and anaphoric dependence: On reconstruction and its implications*. Doctoral dissertation, MIT, Cambridge, Mass.

Belletti, Adriana, and Luigi Rizzi. 1988. Psych-verbs and theta theory. *Natural Language and Linguistic Theory* 6: 291-352.

Boskovic, Zeljko. 1997. *Superiority and economy of derivation: Multiple Wh-fronting*. Ms.

Bresnan, Joan W. 1971. Sentence stress and syntactic transformations. *Language* 47: 257-281.

Campos, Héctor, and Paula Kempchinsky, eds. 1995. *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*. Washington, D. C.: Georgetown University 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. 1986. *Barriers*. 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.

Chomsky, Noam. 1994. Bare phrase structure. MIT occasional papers in linguistics 5. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Also published in Campos and Kempchinsky (1995) and Webelhuth (1995).]

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

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.

den Dikken, Marcel. 1995. Binding, expletives, and levels. *Linguistic Inquiry* 26: 347-354.

Epstein, Samuel D. 1995. *Un-principled syntax and the derivation of syntactic relations*. Ms., Harvard University.

Huang, C. -T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT, Cambridge, Mass.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9: 577-636.

Koizumi, Masatoshi. 1993. Object agreement phrases and the split VP hypothesis. In *Papers on Case and Agreement I: MIT working papers in linguistics* 18, 99-148.

Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Doctoral dissertation, MIT, Cambridge, Mass.

Lasnik, Howard. 1972. *Analyses of negation in English*. Doctoral dissertation, MIT, Cambridge, Mass.

Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2: 1-22.

Lasnik, Howard. 1995a. Last resort and attract F. In *Proceedings of the Sixth Annual Meeting of the Formal Linguistics Society of Mid-America*, ed. Leslie Gabriele, Debra Hardison, and Robert Westmoreland, 62-81. Indiana University, Bloomington, In.

Lasnik, Howard. 1995b. A note on pseudogapping. In *Papers on minimalist syntax, MIT working papers in linguistics* 27, 143-163.

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.

Levin, Nancy. 1978. Some identity-of-sense deletions puzzle me. Do they you? In *Proceedings of the Fourteenth Annual Meeting of the Chicago Linguistic Society*, 229-240. Chicago Linguistic Society, Chicago University, Chicago, Ill.

Lobeck, Anne. 1990. Functional heads as proper governors. In *Proceedings of North Eastern Linguistic Society 20*, 348-362. GLSA, University of Massachusetts, Amherst.

Ochi, Masao. 1997. Move or Attract?: Attract F and the pied-piping chain. Presented at the Open Linguistics Forum, Ottawa.

Postal, Paul M. 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT, Cambridge, Mass. Published as *Infinite syntax!* Norwood, N.J.: Ablex (1986).

Saito, Mamoru, and Keiko Murasugi. 1993. *Subject predication within IP and DP*. Ms., University of Connecticut.

Uriagereka, Juan. In press. Multiple spell-out. In *Working Minimalism*, ed. Marguerite Browning. Cambridge, Mass.: MIT Press.

Webelhuth, Gert, ed. 1995. *Government and binding theory and the minimalist program*. Oxford: Basil Blackwell.

Wyngaerd, Guido Vanden. 1989. Object shift as an A-movement rule. In *MIT working papers in linguistics* 11, 256-271.