**Notes on The Origins and Development of the Cyclic Principle**

I. Origins of the Syntactic Cycle

(1)   The classic theory (Chomsky (1955)): No recursion in the
      base. It is the transformational component that is
      responsible for the infinitude of language. The PS component
      generates only mono-clausal structures. Generalized
      transformations (GTs) embed one structure into another (a
      'matrix'), turning a 'forest' into a 'tree'. Singulary
      transformations (STs) operate on trees only.

(2)   The organization of the syntactic component is as follows:
      Application of the phrase structure rules creates a P-
      marker, or, in the case of a complex sentence, a set of P-
      markers. Then successive application of transformations
      (singularly and generalized) creates successive phrase
      structure representations ('derived P-markers'), culminating
      in a final surface representation. The syntactic levels in
      this theory are that of phrase structure and that of
      transformations, the latter giving a 'history' of the
      transformational derivation (the successive transformational
      steps creating and affecting the structure). The
      representations at these levels are the P-marker and the T-
      marker respectively.

(3)   Fillmore (1963), expanding on a point made by Lees (1963),
      explores the need for limitations ("traffic rules")
      restricting the ways that GTs and STs can interact in a
      derivation.

(4)   If a structure could be embedded into a matrix to which STs
      had already applied, "then the embedding transformations
      would have to become much more complex..." Fillmore (1963)

(5) Chomsky (1965) develops this point in a slightly different
      way: While there are many cases of singulary 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 singulary transformations that must
      apply to a matrix sentence before a sentence transform is
      embedded in it..."

(6) S_1 S_2

       S_2

          S_1
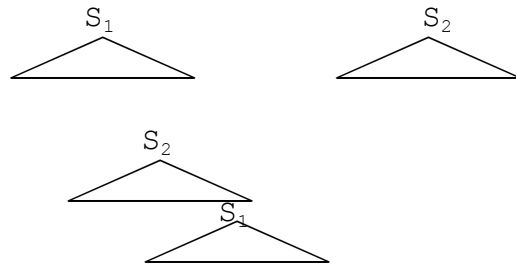
(7) We need to be able to apply STs on S_1, then a GT embedding
     S_1 in S_2, then STs on S_2.
BUT NOT, e.g.,
(8) STs on S_1, then STs on S_2, then a GT embedding S_1 in S_2.

(9) Chomsky gives an additional argument as well:
(10) While there is extensive ordering among singulary
     transformations (situations where a derivation produces an
     unacceptable sentence if two transformations are applied in
     reverse order), "... there are no known cases of ordering
     among generalized transformation although such ordering is
     permitted by the theory of Transformation-markers." Chomsky
     (1965)

(11) Thus, instead of generalized transformations, we should have
     recursion in the base, with transformations applying
     cyclically, first operating on the most deeply embedded
     clause, then the next most deeply embedded, and so on,
     working 'up the tree'.
(12) Thus, singulary transformations invariably apply to
     constituent sentences 'before' they are embedded, and to
     matrix sentences 'after' embedding has taken place.

(13) "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)

(14) Chomsky further claimed that the theory of transformational
     grammar is simplified by this change, since GTs and T-
     markers are eliminated entirely. The P-markers in the
     revised theory (the 'standard theory') contain all of the
     information of those in the LSLT version, but they also
     indicate explicitly how the clauses are embedded in one
     another, that is, information that had been provided by the
     GTs and T-markers.

(15) Curiously, it was almost 3 decades before it became clear
     that the ordering constraints constituted no real argument
     against GTs. It was not recursion in the base that excluded
     (some of) the unwanted derivations: It was the cyclic
     principle. We will return to this issue later in the term.

(16) A cycle-like principle from Chomsky (1965):
(17) The 'Insertion Prohibition': "...no morphological material
     ... can be introduced into a configuration dominated by S
     once the cycle of transformational rules has already
     completed its application to this configuration (though
     items can still be extracted from this constituent of a
     larger 'matrix structure,' in the next cycle of
     transformational rules)."   Chomsky (1965)

(18) "The Insertion Prohibition ... is a step toward a stricter
     interpretation of the cycle: it asserts that once a stage of
     the cycle has been passed, we cannot introduce material into
     it from the outside ..."   Chomsky (1973)

(19) This was proposed to prevent reflexivization into an
     embedded clause (with reflexivization then regarded as a
     rule that deletes some lexical material and inserts other
     lexical material to replace it).
(20) Note that this is weaker (and possibly stronger) than what
     we are used to thinking of as a cyclic constraint.
     a. It allows a movement operation on the S_2 cycle even when
     the moving item and its target are contained in the embedded
     S_1 domain.
     (b. It blocks any lexical insertion into S_1 even if the rule
     involves as a trigger some item in the S_2 domain.)

(21) "To further sharpen the notion 'transformational cycle,'
     suppose that we impose the general condition [(22)]"

(22) The Strict Cycle Condition: No rule can apply to a domain
     dominated by a cyclic node *A* in such a way as to affect
     solely a proper subdomain of *A* dominated by a node *B* which
     is also a cyclic node.

(23) As far as I can tell, (22) simply makes explicit what was
     implicit in the Aspects presentation. If it had not been
     assumed, the unused derivations would not actually have been
     excluded.

(24) "From [(22)] it follows that *wh*-Movement must be a cyclic
     rule, since it applies in indirect questions and relatives."
     Chomsky (1973) [Background: In the 1960's, there was
     extensive discussion of a taxonomy of rules into cyclic and
     post-cyclic or last-cyclic. Chomsky's reasoning indicated
     that post- and last-cyclic rules could apply only in the
     top-most sentence.]

(25) *What did he wonder [where [John put t t]]
(26) Assuming some particular locality constraints (Tensed

Sentence Condition, Specified Subject Condition), to derive
(25), "we must first place *where* in the [Spec of] COMP
position of the embedded sentence. But in that case, *what*
cannot enter the [Spec of] COMP position, which is filled by
where, and thus cannot be extracted on the next cycle. The
principles of the cycle presupposed so far in this
discussion permit no other ordering of rule applications
..."   Chomsky (1973)

(27) McCloskey (1991) shows that Irish has acceptable apparent
long distance WH-movement, but that there is a footprint of
the steps: the Comp takes on a special form. In fact, the
marker shows up even in cases of short movement. (28) is a
complex sentence with no WH-movement. (29) illustrates
relativization, an instance of WH-movement.

(28) Dúirt sé [gur   bhuail tú é]
      said he COMP   struck you him
      'He said that you struck him'
(29) an fear [a   bhuail tú __]
      the man   [COMP struck you ‾‾]
      'The man that you struck'

(30) The marker multiplies in long movement instances, leaving a
clause-by-clause trail from the moved item to its understood
position (and this marking is obligatory, according to
McCloskey):

(31) an rud   a   shíl   mé a   dúirt tú a   dhéanfá
      the thing COMP thought I COMP said you COMP do:COND:2SG]
      'the thing that I thought you said you would do'

II. The Cycle in Phonology

(32) The cyclic principle was first formulated in Chomsky et al.
     (1956). Chomsky and Halle (1968) made crucial use of the
     cycle in virtually every analysis.
(33) The fundamental idea is that complex words have nested
     internal structure, and that (certain) phonological rules
     apply first to the most deeply embedded constituent, then to
     the next most deeply embedded, and so on.
(34) For example, the word stress rule applies in (35) first to
     the individual components black, board, and eraser,
     assigning each a primary stress indicated by '1', then the
     compound stress rule, which reassigns primary stress to the
     left-most primary stress in its domain, applies to black
     board, and finally to black board eraser. Each reassignment
     of primary stress weakens all non-primary stress in the
     domain by one degree.

(35) $[_N$ $[_N$ $[_A$ black] $[_N$ board]] $[_N$ eraser]]
              1   1   1

(36)a $[_N$ $[_A$ black] $[_N$ board]]
                   1   2
    b $[_N$ $[_N$ $[_A$ black] $[_N$ board]] $[_N$ eraser]]
                   1   3   2

(37)"... it is natural to suppose that in general the phonetic
     shape of a complex unit ... will be determined by the
     inherent properties of its parts and the manner in which
     these parts are combined, and that similar rules will apply
     to units of different levels of complexity. These
     observations suggest a general principle for the application
     of rules of the phonological component, namely, what we will
     call the principle of the 'transformational cycle.'" Chomsky
     and Halle (1968), p.15

(38) Alongside the word and compound stress rules, there is a
     rule assigning stress contours to sentences, the Nuclear
     Stress Rule (NSR). Chomsky and Halle propose that this rule
     too is cyclic, in a phonological cycle that follows the
     entire syntactic cycle.
(39) The NSR reassigns primary stress to the right-most primary
     stress in its domain. [Interestingly, the relevant cyclic
     domains for Chomsky and Halle are at least S, NP, and VP, so
     maybe all phrases.]

(40) $[_S$ Mary $[_{VP}$ teaches engineering]]
          1   1   1   word stress
                       2   1 1st cycle NSR
          2   3   1 2nd cycle NSR

(41) Bresnan (1971b) argues that the NSR should not be part of a
     separate cycle, but rather, follow all of the syntactic
     rules on each syntactic cycle, where the cycles are NP, and
     S and/or S' (= modern IP, CP).
(42) She shows how this explains some systematic classes of
     apparent exceptions to the NSR, including one due to Newman
     (1946):
(43) George has plans to leave (plans which he intends to leave)
                          1
Compare:
(44) George has plans to leave (He intends to leave)
                          1

(45) Helen left directions for George to follow
                      1   OR   1

(46)... [$_{NP}$ directions [$_{S'}$ for George to follow directions]]
                 1   1   1   1   word stress
                                2   2   1   Cycle 1 NSR
                                                  Ø   Cycle 2 syntax
                 1   3   3   ?Cycle 2 syntax

(47)... [$_{NP}$ directions [$_{S'}$ for George to follow]]
                 1   1   1   word stress
                                2   1   Cycle 1 NSR
                 2   3   1   Cycle 2 NSR

(48) Bresnan (1971a) analyzes *wanna* contraction in similar terms,
     with the contraction process following the syntactic rules
     on each cycle. (Here, crucially, S and S' are both cyclic
     nodes.)

(49) [$_{S'}$ Q [$_S$ you want [$_{S'}$ for [$_S$ you to visit who]]]
             you want   Ø   Ø   to visit who   EQUI
             you wanna   visit who   Contraction
             _____
         who   you wanna   visit Ø   wh-movement
             do   other rules
             _____
         Who do you wanna visit?
(50) [$_{S'}$ Q [$_S$ you want [$_{S'}$ for [$_S$ who to visit you]]]
                           _____
             _____
         who   you want   Ø to visit you   wh-movement
             do   other rules
             _____
         Who do you want to visit you?

(51) By the time the material intervening between <u>want</u> and <u>to</u> is
     removed, it is too late to do contraction, given strict
     cyclicity, a version of which Bresnan proposes:
(52) "... there is probably a general condition on the
     transformational cycle forbidding a cyclic transformation
     from applying on S_i to effect a structural change entirely
     within S_j if S_i dominates S_j."

III. The Cycle in 'Semantics'

(53) Jackendoff (1969) argues that reflexives are base-generated,
     not the result of transformations, and assigned antecedents
     by an obligatory cyclic rule (the 'reflexive rule').
(54) S and NP are the cyclic nodes.

(55) John showed Bill [a picture of himself]
(56) The reflexive rule does not apply on the first cycle here,
     since there is not even a potential antecedent.
(57) On the second cycle, there are two potential antecedents, so
     we get ambiguity.

(58) *John showed Bill [Mary's picture of himself]
(59) This time, the first cycle does include a potential
     antecedent (a c-commanding Np in more modern terms), so
     coreference is assigned, though anomalously. By the next
     cycle, it is too late to find a more suitable antecedent.

(60) A very theory-internal argument that reflexivization is in
     the transformational cycle, rather than part of a later
     'interpretive' cycle:
(61) *Who did you see stab yourself
(62) you did see [$_S$ who [$_{VP}$ stab yourself]] →
     who did you see [$_{VP}$ stab yourself]
                                (via wh-movement and pruning)
(63) At this point, the reflexive rule should be able to assign
     you as antecedent.
(64) If the rule is in the syntactic cycle, though, on cycle 1,
     who will obligatorily be assigned as antecedent.

(65) Conversely, though I don't think Jackendoff remarks on this,
     in long distance WH-movement cases, we again need to apply
     the reflexive rule before WH-movement:
(66) Who do you think Mary saw stab himself

(67) Lasnik (1972) suggests, on conceptual grounds, that the rule
     determining scope of negation is cyclic.
(68) First, note that it cannot be a deep structure rule, since
     it often depends on the output of transformations (as
     mentioned by Chomsky (1965)):

(69)a Often, I don't attend class often>neg
    b I don't attend class often   neg>often or often>neg

(70) Given strict cyclicity, and given that this rule applies
     uniformly in embedded sentences as well as in main
     sentences, it should be cyclic. [cf. (24) above]

(71) Lasnik (1976) presents an analysis of strong crossover based
     on a cyclic interpretive rule:

(72) *Who does he think can solve the problem [where he is to be
     understood as a variable bound by Who]
(73) "... if the scope of a WH-operator is assigned cyclically
     ... [(72)] will be blocked ... since who will be assigned
     scope on the first possible cycle [who can solve the
     problem], and consequently, he, which is not included in the
     first cyclic domain, will not be bound by who."

(74) Later, Barss (1986) implicitly gave an argument for successive cyclicity of WH-movement based on intermediate reflexive binding (though Barss didn't present this as an argument; rather, he said he assumes successive cyclic movement.

(75) (9) [[which pictures of himself]$_i$ did John think [t'$_i$ [Fred liked t$_i$]]]?

(76) LGB doesn't discuss such examples, as far as I can tell.

(77) Chomsky and Lasnik (1993) discuss them, but, like Barss, assume successive cyclic movement, rather than arguing for it:

(78) "Consider, for example, such constructions as (108), formed by successive-cyclic movement of the question phrase from the position of t, to the position of t', to [Spec, CP] of the matrix clause.
   (108) a. [which picture of himself] did *John* say [t' that
       Bill liked t best]
        b. [which pictures of each other] did *they* say [that
       we liked best]
   Barss (1986) observes that the anaphor can take either of the italicized NPs as its antecedent. But an anaphor can only be bound by the closest c-commanding subject, as we see in the corresponding expressions (109), without wh-movement.
   (109) a. John said [that Bill liked [that picture of
       himself] best]
        b. they said [that we liked [those pictures of each
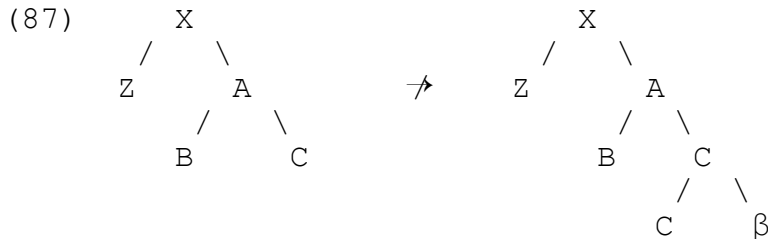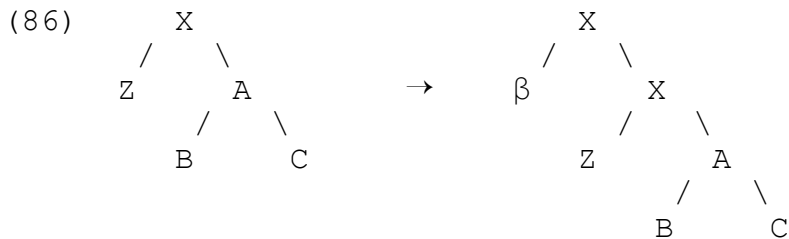   other] best]
   Here the antecedents must be *Bill*, *we.* In (108) the same binding condition requires that each of the traces be "visible," the question phrase being interpreted for binding as if it were in one or the other of these positions (*chain binding*)."

IV. Further Syntactic Developments

(79) "... transformational rules, e.g., [Move NP and Move *wh-phrase*], meet the condition of the (strict) cycle ..."
   Chomsky (1977, p.73)

(80) Chomsky (1993), in the course of his minimalist critique of all of syntactic theory, argues that deep structure should be eliminated and generalized transformations reinstated. He notes that in Aspects "Elimination of generalized transformations in favor of cyclic base generation [sic] is ... justified in terms of explanatory adequacy. But the questions under discussion then do not arise in the far more restrictive current theories."   Chomsky (1993)

(81) We then have 'bottom-up' creation of a phrase marker by
     generalized transformations. There is no 'base'.
(82) The derivation is constrained by an 'extension condition':
     "... GT and Move α extends K to K*, which includes K as a
     proper part." Chomsky (1993)
(83) "...[(82)] yields a version of the strict cycle..."

(84) One further consequence? "...the binarity of GT comes close
     to entailing that X-bar structures are restricted to binary
     branching ..." p.191 [Note, though, the presupposition.]

(85) The Extension Condition: "operations preserve existing
     structure." Chomsky (2000)

```
(86)      X                        X
        /   \                    /   \
      Z       A       →        β       X
            /   \                    /   \
          B       C                Z       A
                                         /   \
                                       B       C


(87)      X                        X
        /   \                    /   \
      Z       A       ↛        Z       A
            /   \                    /   \
          B       C                B       C
                                         /   \
                                       C       β
```

(88) The original tree is a sub-tree of the derived tree in
     derivation (86) but not in derivation (87).

(89) {X, ZA, ZBC}   →   {X, βX, βZA, βZBC}

(90) {X, ZA, ZBC}   →   {X, ZA, ZBC, ZBCβ}

(91) Surprisingly, though, with the **'low'** attachment, the
     original structure is a subset of the derived one, while
     with the 'high' attachment, it is not.
(92) More generally, the higher the attachment site, the more
     radically the set is altered.

V. Successive cyclicity and A-movement

(93) John seems [to be certain [to win]]
(94) "If the rule of NP-movement that yields [(93)] applies
     successive cyclically ... then the rule will observe
     subjacency." Chomsky (1977)

(95) John was believed to have been killed
(96) John INFL be believed [$_S$ t' INFL have been killed t]

(97) "In the case of sentence [(96)], we are led by the
     projection principle [sic] to assume that the rule Move-α
     applies twice, leaving the two traces *t* and *t'*,
     successively."   Chomsky (1981)

(98) they are likely [t' to appear to each other [t to be happy]]

(99) "The GF-$\bar{\theta}$ filled by medial traces such as *t'* in [(96)] may
     ... be relevant to LF; for example in the sentence [(98)],
     ... where the medial trace serves as the antecedent of *each
     other*, which requires an antecedent in the same clause in
     such sentences in accordance with binding theory ..."
     Chomsky (1981)

(100) It actually isn't completely clear that the antecedent must
      be in the same clause, given the binding theory in Chomsky
      (1981), or those in Chomsky (1973) and Chomsky (1986) for
      that matter.

(101) And there is an argument in the literature that A-movement
      is **not** successive cyclic.

(102) Epstein and Seely (1999), Epstein and Seely (2006) present
      the following argument: successive cyclic A-movement
      creates a chain. According to Chomsky (1995), a chain is a
      set of 'occurrences' where each occurrence is defined in
      terms of sisterhood. Since each intermediate trace position
      is a Spec of some X, its sister is X', an intermediate
      projection of X. But it is widely assumed that syntactic
      operations can't target intermediate projections. Therefore
      the needed chain links can't exist.

(103) Possible responses:
(104) Is it completely clear that syntactic operations can't
      target X'? I actually believe that the assumption is
      correct, but it is interesting to note that very little
      actual evidence has been offered in the literature.
(105) Given bottom-up GT derivations, at the point where the
      intermediate positions will be created, the moving DP will
      be targeting a <u>maximal</u> projection.

(106) And there is some evidence for a clause-mate requirement
      (stronger than the Tensed Sentence Condition of Chomsky
      (1973) or the Governing Category requirement of Chomsky
      (1981) and Chomsky (1986)). [See also Postal (1966), Postal
      (1974).]

(107) Condition A [from Lasnik (2002)]:

(108)a Jack made himself out to be immoral
    b ?*Jack made out himself to be immoral
(109)a They made each other out to be honest
    b ?*They made out each other to be honest

(110) ?Jack called up himself
(111) ?They called up each other

(112) John appears to Mary [to seem to himself/*herself [to
    be the best candidate]]   [pointed out to me in this
    connection by Adolfo Ausín; also attributed to Danny Fox,
    via David Pesetsky, in Castillo et al. (1999)]

(113) Condition B:

(114) *John$_i$ injured him$_i$
(115) *John$_i$ believes him$_i$ to be a genius

(116) *Mary injured him$_i$ and John$_i$ did too
(117) ?Mary believes him$_i$ to be a genius and John$_i$ does too

(118) Weak pronouns must cliticize onto the verb. Oehrle
    (1976)
(119) The detective brought him in
(120) *The detective brought in him   Chomsky (1955)

(121) Failure to cliticize in (117) is repaired by ellipsis.
(122) In (116), on the other hand, the pronoun and its antecedent
    are clause-mates independent of cliticization.

(123) *John$_i$ is believed [to seem to him$_i$ [to be a genius]]

References

Barss, Andrew. 1986. *Chains and anaphoric dependence: On
    reconstruction and its implications*. Doctoral dissertation,
    MIT, Cambridge, Mass.
Bresnan, Joan W. 1971a. Contraction and the transformational
    cycle. Ms. MIT, Cambridge, Mass.
Bresnan, Joan W. 1971b. Sentence stress and syntactic
    transformations. *Language* 47: 257-281.
Castillo, Juan Carlos, John Drury and Kleanthes K. Grohmann.
    1999. Merge over Move and the Extended Projection Principle.
    In *University of Maryland Working Papers in Linguistics* 8,
    66-104. (Mistakenly listed in table of contents as "The
    Status of the Merge over Move Preference.)
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. 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. 1977. On *wh*-movement. In *Formal syntax*, ed. Peter
    Culicover, Thomas Wasow, and Adrian Akmajian, 71-132. New
    York: Academic Press.
Chomsky, Noam. 1981. *Lectures on government and binding*.
    Dordrecht: Foris.
Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.
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 and Morris Halle. 1968. *The sound pattern of
    English*. New York: Harper and Row.
Chomsky, Noam, Morris Halle and Fred Lukoff. 1956. On accent and
    juncture in English. In *For Roman Jakobson*, ed. M. Halle, H.
    Lunt, and H. MacLean, 65-80. The Hague: Mouton & Co.
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. 1999. SPEC-ifying the GF
    "subject"; eliminating A-chains and the EPP within a
    derivational model. Ms. University of Michigan and Eastern
    Michigan State University,
Epstein, Samuel D. and T. Daniel Seely. 2006. *Derivations in
    Minimalism*. Cambridge: Cambridge University Press.
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.
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. 2002. Clause-mate conditions revisited. *Glot International* 6: 94-96.

Lees, Robert B. 1963. *The grammar of English nominalizations*. The Hague: Mouton. (Originally published 1960 byIndiana University. Research Center in Anthropology, Folklore, and Linguistics.)

McCloskey, James. 1991. Resumptive pronouns, A'-binding, and levels of representation in Irish. In *Syntax and Semantics 23: The Syntax of the Modern Celtic Languages*, ed. Randall Hendrick, 199-248. New York: Academic Press.

Newman, Stanley. 1946. On the stress system of English. *Word* 2: 171-187.

Oehrle, Richard. 1976. *The grammatical status of the English dative alternation*. Doctoral dissertation, MIT, Cambridge, Mass.

Postal, Paul M. 1966. A note on 'understood transitively'. *International Journal of American Linguistics* 32: 90-93.

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