Analyzing the English Aux: An Historical Perspective

Howard Lasnik
University of Connecticut

I. Syntactic Structures (1957)

(1) John left           John didn't leave
John should leave       John shouldn't leave
John has left           John hasn't left
John is leaving         John isn't leaving

(2) *John leftn't       *John didn't should leave
*John doesn't have left
*John doesn't be leaving

(3) John left           Did John leave
John should leave       Should John leave
John has left           Has John left
John is leaving         Is John leaving

(4) *Left John          *Did John should leave
*Does John have left    *Does John be leaving

(5) S
    NP  VP
       | Verb
    Aux  V

(6) Aux = C (Modal) (have en) (be ing)

(7) C = 0 in other contexts
    past in any context

(8) T\textsubscript{unt} - optional \#16
    NP \rightarrow C \rightarrow V...  
    NP \rightarrow C+M \rightarrow ...
    Structural analysis:  NP \rightarrow C+have \rightarrow ...
    Structural change:  X\_1 \rightarrow X\_2 \rightarrow X\_1 \rightarrow X\_2 \rightarrow A \rightarrow X\_1

(9) T\textsubscript{a} - optional \#17
    Structural analysis: same as \#16
    Structural change:  X\_1 \rightarrow X\_2 \rightarrow X\_1 \rightarrow X\_2 \rightarrow A \rightarrow X\_1

(10) T\textsubscript{c} - optional \#18
    Structural analysis: same as \#16
    Structural change:  X\_1 \rightarrow X\_2 \rightarrow X\_1 \rightarrow X\_2 \rightarrow X\_1

(11) Auxiliary Transformation ("Affix Hopping") - obligatory \#20
    Structural analysis:  X - Af - v - Y (where Af is any C or \textit{is} en or \textit{ing}; v is any M or V, \textit{or} have or be)
    Structural change:  X\_1 \rightarrow X\_2 \rightarrow X\_1 \rightarrow X\_1 \rightarrow X\_2 \rightarrow X\_1

(12) Word Boundary Transformation - obligatory \#21
    Structural analysis:  X - Y (where X\#v or Y \#Af)
    Structural change:  X\_1 \rightarrow X\_2 \rightarrow X\_1 \rightarrow X\_2

(13) do - Transformation - obligatory \#22
    Structural analysis:  # - Af
    Structural change:  X\_1 \rightarrow X\_3 \rightarrow X\_2 \rightarrow X\_1

(14) The fundamental insight of this system is that the tense-agreement inflectional morpheme ("C") is syntactically independent, even though always a bound morpheme superficially. The analysis is brilliantly successful, but when viewed from the perspective of explanation in the sense of Chomsky (1965), it has serious shortcomings.

II. Verb Raising Analyses

(15) In the base, Aux includes only C and, optionally, Modal. When there is no modal, the 1st instance of have or be following the Aux is raised into the Aux. This makes possible a substantial limitation on the descriptive power of transformations: a non-variable term must be a constituent. The non-constituent terms in (8)-(10) above become simply Aux in such an analysis.

(16)a have-be Raising - obligatory
b Affix Hopping - demands adjacency between Af and v - obligatory
c do-support - obligatory and strictly ordered after a.

(17) Restatement in terms of 'head movement':
a S is the maximal projection of the inflectional morpheme Infl (= C of Syntactic Structures).  
b Infl takes VP as its complement.
c When the head of VP is have or be it raises to Infl, the next head up. (\textit{not} is a modifier of VP?)
d Otherwise Infl lowers to V (under a condition of adjacency?).
e Otherwise do adjoins to Infl.

(18) The 'stranded affix' filter: A morphologically realized affix must be a syntactic dependent of a morphologically realized category, at surface structure. (Lasnik (1981))
(19) UG principles are applied wherever possible, with language-particular rules used only to "save" a D-structure representation yielding no output. Verb raising and affix hopping are universal; do-support is language-particular. (Chomsky (1991))

(20) a *John likes not Mary
b Jean (n’)aime pas Marie

(21) In French, all verbs are capable of raising, not just have and be. Unlike the situation in English, affix hopping and do-support are never needed. (Emonds (1978))

(22) 'Infl' is not one head; it consists of (at least) Tense and Agr, each heading its own projection.

(23) a English Agr, because not morphologically rich, is 'opaque' to θ-role transmission. Thus, if a verb with θ-roles to assign were to raise, it would be unable to assign them, resulting in a violation of the θ-criterion.
b French Agr, because morphologically rich, is 'transparent' to θ-role transmission. (Pollock (1989))

III. Economy of Derivation

(24) Raising is preferred to lowering, because lowering will leave an unbound trace that will have to be remedied by re-raising in LF. (Chomsky (1991))

(25) a *John not writes books
b John does not write books

(26) Why isn't (25)a, with overt affix lowering followed by LF 're-raising', preferred over (25)b, with language particular last resort do-support?

(27) The Head Movement Constraint (reduced to an ECP antecedent government requirement) prevents the LF re-raising needed in the derivation of (25)a. The intervening head NEG cannot be crossed.

(29) But then why is overt raising possible in French, and, in the case of have and be, in English as well?

(30) a If AGR moves, its trace can be deleted, since it plays no role in LF. [[(We actually need something stronger: The trace must be deleted, and must be deleted immediately.]]
b If V moves, its trace cannot be deleted.
c Deletion of an element leaves a category lacking features, [e].
d Adjunction to [e] is not permitted. (Chomsky (1991))

(31) a When V overtly raises, (20)b, it first adjoins to AGR, creating \{AGR, V AGR\};
b Next, AGR raises to T, crossing NEG, thus leaving a trace that is marked [−γ], indicating a violation of the ECP. That trace is an AGR;
c Eventually, in accord with (30)a, the [−γ] trace is deleted, so there is no ECP violation (where ECP is, as in Lasnik and Saito (1984;1992), an LF filter: *[−γ].

(32) a When V vainly attempts to covertly (re-)raise in LF, (25)a, AGR has already lowered overtly to T, leaving an AGR trace (which deletes, leaving [e]), and creating a complex T,
b which has lowered to AGR, leaving a T trace and creating a still more complex AGR,
c which has lowered to V, leaving an AGR trace (which deletes, leaving [e]), and creating a complex V.
d This complex V raises to the [e] left by the deletion of the AGR trace, a movement that is, by (30)d, necessarily substitution, thus turning [e] into V.
e This element now raises across NEG to (the trace of) T, leaving behind a [−γ] trace which is, crucially, a V trace, hence non-deletable. The resulting LF is in violation of the ECP.

(33) Note that (30)a, (31)c might be inconsistent with a central economy condition of Chomsky (1991): Deletion is only permitted to turn an ill-formed LF object onto a well-formed LF object, where the relevant well-formed objects are 'uniform chains' (chains all of whose members are X0s, are in A-positions, or are in A'-positions. This is precisely to prevent making a short licit head-, A-, or adjunct-movement, followed by a long illicit movement, with subsequent deletion of the offending trace. But exactly that is crucially being allowed here.

(34) Another possible problem is that generally, an illicit movement results in some degradation (e.g., Subjacency effects), even if the offending trace is eventually eliminated. But the overt V-movement at issue here is fully grammatical.
IV. A Minimalist Approach (Chomsky (1993))

(35)a Strong lexicalism: verbs are pulled from the lexicon fully inflected.
b There is thus no need for affix hopping. 
c Rather, the inflected V raises to Agr (and T) to 'check' the features it already has. This checking can, in principle, take place anywhere in a derivation on the path to LF. 
d Once a feature of Agr has done its checking work, the feature disappears.

(36) So what's the difference between French and English?

(37)a In French, the V-features of Agr (i.e., those that check features of a V) are 'strong'. 
b In English, the V-features of Agr are 'weak'.

(38)a If V raises to Agr overtly, the V-features of Agr check the features of the V and disappear. If V delays raising until LF, the V-features of Agr survive into PF. 
b V-features are not legitimate PF objects. 
c Strong features are visible at PF; weak features are not. Surviving strong features cause the derivation to 'crash' at PF.

(39) This forces overt V-raising in French.

(40) In English, delaying the raising until LF does not result in an ill-formed PF object, so such a derivation is possible. What makes it necessary is:

(41) 'Procrastinate'? Whenever possible, delay an operation until LF.

(42) Why do have and be raise overtly?

(43) Have and be are semantically vacuous, hence not visible to LF operations. Thus, if they have not raised overtly, they will not be able to raise at all. Their unchecked features will cause the LF to crash.

(44) Questions about (43): (1) Should syntactic operations, even those in the LF component, care about purely semantic properties? (2) There are languages (such as Swedish in (45)) where auxiliary verbs have inflectional features but do not raise overtly. (3) Even instances of have and be arguably possessing semantic content raise overtly.

(45)a ..., om hon inte ofte har sett honom whether she not often has seen him 
b * om hon har inte ofte sett honom 
c * Om hon inte har ofta sett honom

(32)a Is there a solution / There isn't a solution 
b Have you any money / I haven't any money

(46) *John not left
(47) Chomsky (1993) does not discuss how to rule out (46). Note that (32) does not carry over to this framework (even if we wanted it too), since (32) crucially relies on affix hopping.

V. Notes Towards a Hybrid Minimalist Account

(48) Chomsky's minimalist account demands that Agr and T are just abstract features that check against features of verbs when verbs raise to them. All the earlier accounts treated such Inf items as bound morphemes that had to become affixes. Can both possibilities coexist?

(49)a French verbs are fully inflected in the lexicon (possibly correlating with the fact that there are no bare forms; even the infinitive has an ending). 
b Have and be are fully inflected in the lexicon (possibly correlating with the fact that they are highly suppletive). 
c All other English verbs are bare in the lexicon.

(50) Inf is freely an affix or a set of abstract features.

(51)a Featural Inf is strong in both languages. 
b Affixal Inf must merge with a V, a PF process demanding adjacency. 
(52)a ... Inf ... V ... OK. V will overtly raise. +F +F 
b ... Inf ... V ... OK. PF merger. Af bare 
c ... Inf ... V ... * at LF. +F of I won't be checked. +F bare 
d ... Inf ... V ... * at LF. +F of V won't be checked. Af +F (Maybe * at PF also, if merger fails.

(53)a French Inf will thus always have to be featural. 
b English Inf will always have to be featural, when the verb is have or be. 
c English Inf will always have to be affixal with any other verb.

(54)a *John not left {Merger couldn't have taken place.} 
b *John left not {Left isn't in the lexicon, so no feature could drive raising.}

(55) Jean (n')aime pas Marie

(56) John has not left

(57) Why is raising allowed in (55), (56)? Here are 3 possibilities:

(58)a NEG is not a head, but a modifier. Note that its major role as a head had been to block (54)a, which is now irrelevant to the issue.
b The Head Movement Constraint is 'relativized' to different kinds of heads, as in Roberts (1994).
c {The most radical} There is no Head Movement Constraint. In any theory where movement is driven solely by the need for features to be satisfied, the standard HMC example is irrelevant: *Read John will the book won't be generated simply because no feature will drive the movement of read to Comp. It is only finite verbs that raise to Comp, indicating that the crucial feature is Tense.

VI. A Surprising Paradigm: Evidence for the Hybrid?

(59) John slept, and Mary will too
(60) a *John slept, and Mary will slept too
   b John slept, and Mary will sleep too

(61) ?John was sleeping, and Mary will too
(62) a *John was sleeping, and Mary will sleeping too
   b John was sleeping, and Mary will sleep too

(63) John has slept, and Mary will too
(64) a *John has slept, and Mary will slept too
   b John has slept, and Mary will sleep too

(65) Hypothesis 1: Any form of a verb V can be 'deleted under identity' with any form of V (reminiscent of Fiengo and May's 'vehicle change').

(66) *John was here, and Mary will too
(67) a *John was here and Mary will was here too
   b John was here and Mary will be here too

(68) Could it be that a trace can't serve as (part of) an antecedent for deletion?
(69) Linguistics, I like t, and you should like linguistics too
(70) ?Someone will be t in the office, won't there be someone in the office?
(71) That this approach will fail is likely t. Yes it is likely that this approach will fail.

(72) Hypothesis 2: A form of a verb V can only be deleted under identity with the very same form. Forms of be and auxiliary have are introduced into syntactic structures already fully inflected. Forms of 'main' verbs are created out of lexically introduced bare forms and independent affixes.

(73) a John is not foolish
   b *Be not foolish
   c Be foolish

(74) a The Imperative morpheme (generated in the position of Tense) is strictly affixal, hence there will never be raising to it (just merger with it)
   b OR Imp is freely affixal or featural, and be and auxiliary have are defective, lacking imperative forms in the lexicon.

(75) a *Not leave (Lack of adjacency blocks merger)
    b *Not be foolish

(76) Leave. I don't want to.
(77) Mary left. I don't want to.
(78) Be quiet. I don't want to.
(79) Mary is quiet. *I don't want to.

VII. An Alternative Treatment of the Gap in the Paradigm?

(80) " [v [ e ] X ] cannot antecede VP-ellipsis." Roberts (n.d.)
(81) "...a trace of verb movement cannot serve as part of a VPE antecedent." Potsdam (1996)
(82) A number of languages with overt V raising to I nonetheless allow VP ellipsis, with the effect that everything in the VP except the V is deleted. Doron (1990) shows this for Hebrew:
(83) Q: Salaxt et ha-yeladim le-beit ha-sefer
   A: Salaxti
   "Did you send the kids to school?"
   "I did"

(84) Martins (1994) shows the same thing for Portuguese and McCloskey (1990) does for Irish:
(85) A Martas deu um livro ao João? Sim, deu.
   The Martha gave a book to the John yes gave
   "Did Martha give a book to John? Yes, she did."
(86) Q: Ar chuir tú isteach air
   A: Chuir put [PAST] you in on it
   "Did you apply for it?"
   "Yes."

(87) [v [ e ] X ] cannot antecede VP-ellipsis of [v [ V ] X ].
(88) "...a raised V has fewer features than a non-raised V, assuming that the features that cause raising are not copied (this has to be assumed in a minimalist framework or the raising operation would not eliminate features and so would have no motivation, and so would be impossible given the general last-resort nature of movement)." Roberts (n.d.)

VIII. Problems for the Alternative

(89) A candidate for a VP headed by verb trace anteceding deletion of a VP headed by a lexical verb: Pseudogapping as overt NP raising to Spec of AgrO followed by VP ellipsis. (Lasnik (1995a), based on the proposal of Koizumi (1993), following Johnson (1991), that 'object shift' is overt in English)
a. John hired Bill and Mary will Susan
b. John [VP hired [AgrP Bill [VP t t]]] and Mary will [AgrP Susan [VP hire t]]

Crucially, Pseudogapping is not just deletion of the verb:

The DA proved Jones guilty and the Assistant DA will prove Smith guilty

If (87) is correct, it should presumably generalize to all heads, not be limited to V and trace of V:

[VP t e X ] cannot antecede YP-ellipsis of [VP Y ] X .

Sluicing (Ross (1969)), now standardly analyzed as IP ellipsis (Lobeck (1990) and Saito and Murasugi (1990)), provides another potential counter-example.

Speaker A: Mary will see someone
Speaker B: Tell me who Mary will see
Speaker A: Mary will see someone
Speaker B: Who Mary will see
Speaker A: Never will [IP Harry t go to a linguistics lecture again]
Speaker B: Tell me why [IP Harry will never go to a linguistics lecture again]
Speaker A: Never will [IP Harry t go to a linguistics lecture again]
Speaker B: Why [IP Harry will never go to a linguistics lecture again]

IX. Why Isn't Roberts' Line of Reasoning Valid?

Given that a raised X^0 has had a feature (or set of features) checked and deleted, why can it antecede the deletion of an XP with its head in situ (as in Pseudogapping and Sluicing)?

An ultimately related question: Given that NP raises but V doesn't raise in the Pseudogapping construction, why must V raise in corresponding non-elliptical version?

*Mary will Susan hire

A parallel question: Given that Infl doesn't raise to Comp in the Sluicing construction, why must Infl raise in the corresponding matrix non-elliptical version?

Which linguists Susan will never understand

Overt movement is driven by a 'strong feature' of a head, which attracts a matching feature within the complement of that head. All movement, whether covert or overt, is fundamentally feature movement. [Chomsky (1995b)]

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. Chomsky (1995a, p.262)

Applied to the feature F, the operation Move thus creates at least one and perhaps two "derivative chains" alongside the chain CHF=(F,tF) constructed by the operation itself. One is CHFF=(FF[FF],tFF[FF]), consisting of the set of formal features FF[FF] and its trace; the other is CHCAT=(α,τα), α a category carried along by generalized pied-piping and including at least the lexical item containing F. CHFF is always constructed, CHCAT only when required for convergence... As noted, CHCAT should be completely dispensable, were it not for the need to accommodate to the sensorimotor apparatus." [p.265]
morphology and the internal structure of phrases. Note that such considerations could permit raising without pied-piping even overtly, depending on morphological structure..."

[p.264]

(113) In (109), if only the attracted features raise, but the V does not raise, a PF crash will ensue, but only if the offending item exists at that level. Deletion provides another way to salvage the derivation. When the lower VP is deleted without the V having raised, a PF crash is avoided and the result is acceptable Pseudogapping.

(114) CP

/ \  
NP C'  
who / \  
C IP  
[strong F] / \  
NP I'  
Mary / \  
Infl VP  
will |  
[F] V'  
/ \  
V NP  
see t

(115) An account completely parallel to that provided for (109) is available for (114).

(116) Note that now, the major prima facie counter-examples to Roberts' proposal ((87), as generalized to (95)), are completely compatible with it.

(117) So why not accept the Roberts-Potsdam account of the gap in the original ellipsis paradigm?

(118) John slept, and Mary will too

(119) *John was here, and Mary will too

(120) John was here, and Mary will be here too

(121) Here be does not raise at all, with or without pied-piping, whereas was obviously does raise, resulting in features being checked and deleted.

(122) What are those features? It is hard to see how they could be anything other than inflectional features. But checking and deleting the inflectional features of was makes it more like be, not less like be.

X. Another Kind of Justification for (95)

(123) [Under ellipsis] Corresponding X traces [unlike XP traces] must have the same binder in both the antecedent and target clauses.

(124) Chicken, she'll eat, but ostrich, she won't

(125) Potsdam claims that in Hebrew and Irish, both V-raising languages that have VP ellipsis, "the raised verbs in ellipsis antecedent and target clauses must be the same." He suggests that (123) is universal.

(126) Q: dina soret et ha- svederim Se- hi loveSet Dina knits ACC the sweaters that she wears "Does Dina knit the sweaters that she wears?"

A1: lo, aval ima Sela soretet no, but mother hers knits "No, but her mother does."

A2: lo, ima Sela kona (la) no, mother hers buys (to-her) "No, her mother buys them (for her)." Hebrew

Doron (1990)

(127) A1 is 'strict' or 'sloppy'. A2 is only strict.

(128) Ivan piše rad pažljivo, njegov asistent čita Ivan writes paper carefully and his assistant is reading it carefully.

(129) Marko gradi se kucu, a Marija kupuje Marko builds himself house and Marija buys "Marko is building himself a house, and Maria is buying herself a house."

(130) Q: Does Dina knit the sweaters that she wears? A: No her mother, buys the sweaters that she, wears

(131) The putative answer (130)A is strikingly unresponsive to the question.

(132) dina soret et ha-svederim Se- hi loveSet, Dina knits the sweaters that she wears be-?od ima Sela kona while mother hers buys

(133) dina soret et ha-svederim Se- hi loveSet, Dina knits the sweaters that she wears be-?od ima Sela kona otam while mother hers buys them

(134) Dina knits the sweaters that she wears while her mother buys them

(135) dina ohevet ko sveder Se- hi loveSet Dina loves every sweater that she wears aval ima Sela sonet otam but mother hers hates them "Dina loves every sweater that she wears but her mother hates
them."

(136)  dina ohevet ko sveder Se- hi loveSet
Dina loves every sweater that she wears
av al ima Sela sonet
but mother hers hates
"Dina loves every sweater that she wears but her mother hates
every sweater that she
wears."

Bibliography


Watanabe, Akira. 1993. AGR-based Case theory and its interaction with the A-bar system. Cambridge,
Movement, ed. by David Lightfoot and Norbert Hornstein, 305-50. Cambridge: Cambridge University
Press.
Sprachwissenschaft in Frankfurt Arbeitspapier 10. Frankfurt.