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  John did leave 
    John should leave  John should leave 
    John has left  John has left 
    John is leaving  John is leaving 

(4) (*John left  
    *John did should leave  
    *John does have left  
    *John does be leaving 

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

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

(7) \[ \begin{array}{c} 
  S \\
  NP \rightarrow C (Modal) (have en) (be ing) \\
  \end{array} \]

(8) Aux → C (Modal) (have en) (be ing) 

(9) C \rightarrow \begin{cases} 
  S \text{ in the context } NP_{\text{sing}} \\
  \varnothing \text{ in other contexts} \\
  \text{past in any context} 
\end{cases} 

II. Verb Raising Analyses

(10) T_{\text{not}} \rightarrow \text{optional} \quad \#16 
    \begin{cases} 
      \text{NP} \rightarrow C \rightarrow V ... \\
      \text{NP} \rightarrow C \rightarrow M ... \\
    \end{cases} 
    \text{Structural analysis:} \\
    \text{Structural change: } X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_1 \rightarrow X_2 + A \rightarrow X_3 

(11) T_{a} \rightarrow \text{optional} \quad \#17 
    \text{Structural analysis: same as #16} 
    \text{Structural change: } X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_2 - X_1 - X_3 

(12) T_{a} \rightarrow \text{optional} \quad \#18 
    \text{Structural analysis: same as #16} 
    \text{Structural change: } X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_2 - X_1 - X_3 

(13) Auxiliary Transformation \rightarrow \text{obligatory} \quad \#20 
    \text{Structural analysis: } X \rightarrow A_f \rightarrow V \rightarrow Y \text{ (where } A_f \text{ is any } C \text{ or is } en \text{ or } ing; \text{ or any } M \text{ or } V, \text{ or have or be}) 
    \text{Structural change: } X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4 \rightarrow X_2 - X_3 - X_2 # - X_4 

(14) Word Boundary Transformation \rightarrow \text{obligatory} \quad \#21 
    \text{Structural analysis: } X \rightarrow Y \text{ (where } X \neq V \text{ or } Y \neq A_f) 
    \text{Structural change: } X_1 \rightarrow X_2 \rightarrow X_1 - X_2 

(15) do \rightarrow \text{Transformation} \rightarrow \text{obligatory} \quad \#22 
    \text{Structural analysis: } # \rightarrow A_f 
    \text{Structural change: } X_1 \rightarrow X_2 \rightarrow X_1 - \text{do} + X_2 

(16) 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. Further, it said nothing about uncontracted not, or negation in imperative sentences.
(19) 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. (not is a modifier of VP?)
   d  Otherwise Infl lowers to V (under a condition of
       adjacency?).
   e  Otherwise do adjoins to Infl.

(20) The 'stranded affix' filter: A morphologically realized
     affix must be a syntactic dependent of a morphologically
     realized category, at surface structure. (Lasnik (1981))

(21) (20) eliminates much of the strict rule ordering and
     arbitrary obligatory marking, but does not guarantee that
     do-support is a 'last resort', operating only when there is
     no other way to avoid a stranded affix (see (2), (4), (6)
     above).

(22) A syntactic version of the 'Elsewhere Condition' of
     Kiparsky (1973): If transformations T and T' are both
     applicable to a P-marker P, and if the set of structures
     meeting the structural description of T is a proper subset
     of the set of structures meeting the structural description
     of T', then T' may not apply. (Lasnik (1981))

(23) In the terms of (18), the SDs of verb raising and affix
     hopping mention Infl and (aux) V, while that of do-support
     mentions only Infl.

(24) 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))

(25)a  *John likes not Mary
     b  Jean (n')aime pas Marie

(26) 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. (Beesna (1978))

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

(28)a  English Agr, because not morphologically rich, is 'opaque'
     to 0-role transmission. Thus, if a verb with 0-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 0-role transmission. (Pollock (1989))

III. Economy of Derivation

(29) 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))

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

(31) Why isn't (30)a, with overt affix lowering followed by LF
     re-raising, preferred over (30)b, with language particular
     last resort do-support?

(32)

(33) The Head Movement Constraint (reduced to an ECP antecedent
     government requirement) prevents the LF re-raising needed in
     the derivation of (30)a. The intervening head NEG cannot be
     crossed.

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

(35)a  If AGR moves, its trace can be deleted, since it plays no
     role in LF.
     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))

(36)a  When V overtly raises, (25)b, it first adjoins to AGRo,
     creating [AGRO V AGRo];
     b  Next, AGRo 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 (35)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].)
When V vainly attempts to covertly (re-)raise in LF, AG~5 has already lowered overtly to T, leaving an AGR trace (which deletes, leaving [e]), and creating a complex T.

This complex V raises to the [e] left by the deletion of the AGR trace, a movement that is, by (35)d, necessarily substitution, thus turning [e] into V.

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.

Note that (35)a, (36)c are 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 X's, 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.

Another 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

A. (Chomsky (1993))

Strong lexicalism: verbs are pulled from the lexicon fully inflected.

There is thus no obvious need for affix hopping.

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.

Once a feature of Agr has done its checking work, it disappears.

So what's the difference between French and English?

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

In English, the V-features of Agr are weak.

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.

This forces overt V-raising in French.

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:

'Procrastinate': Whenever possible, delay an operation until LF.

Why do have and be raise overtly?

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.

*John not left

Chomsky (1993) does no discuss this problem. I assume he had in mind some version of (37) (with all its difficulties).

B.

At the core of 'economy' approaches, of which the 'minimalist' approach is one, is the concept of choosing the best among competing derivations. It has never been clear in general, however, what determines the relevant comparison set.

Chomsky has recently suggested a highly principled answer: To begin a derivation, you choose from the lexicon all the items you will use, annotating each with a counter indicating how many times it will be used. Call this collection a 'numeration'. The comparison set includes all and only derivations from the same numeration. This has the positive effect that (52)a does not block (52)b (or vice versa), since the numerations differ with respect to there.

There is someone here

Someone is here

In line with strong lexicalism, forms of do, just as much as there, are in the lexicon. Do, when it occurs, will then be part of a numeration. Derivations with and without do are not comparable. The 'last resort' nature of do-support cannot be directly captured.

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 Infl items as bound morphemes that had to become affixes. Can both possibilities coexist?
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).

- **Have** and **be** are fully inflected in the lexicon (possibly correlating with the fact that they are highly suppletive).
- All other English verbs are bare in the lexicon.

**Infl is freely an affix or a set of abstract features.**

- **Featural Infl** is always strong (as are possibly all featural functional heads).
- **Affixal Infl** must merge with a V, a PF process demanding adjacency. (Halle and Marantz (1993); Bobaljik (1993))

- Infl v OK. V will overtly raise.
- Infl... v OK. PF merger.
- Infl... v * at LF. +F of I won't be checked.
- Infl... v * at LF. +F of V won't be checked.

French Infl will thus always have to be featural.

- English Infl will always have to be affixal, when the verb is have or be.
- English Infl will always have to be affixal with any other verb.

*John not left {Left isn’t in the lexicon; merger couldn’t have taken place.}

- *John left not {Left isn’t in the lexicon, so no feature could drive raising.}

Jean (n’)aime pas Marie

John has not left

Why is raising allowed in (61), (62)? Here are 3 possibilities:

- (37) above, as in Chomsky (1991).
- NEG is not a head, but a modifier. Note that its major role as a head had been to block (60)a, which is now irrelevant to the issue.
- (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, clearly indicating that the crucial feature is Tense.

- John is not foolish
- *Be not foolish
- Be foolish

The Imperative morpheme (generated in the position of Tense) is strictly affixal, hence there will never be raising to it (just merger with it)

- OR Imp is freely affixal or featural, and be and auxiliary have are defective, lacking imperative forms in the lexicon.

- *Not leave {Lack of adjacency blocks merger}
- *Not be foolish

The first (problem) shall be the last: What determines the distribution of ‘supportive’ do?

Conjecture: do doesn’t support; rather it needs to be supported, in the fashion of the LF affix pleonastic there. Case makes there visible as a target of movement. What makes do visible?

- [+Q] Did John leave?
- [+Neg] John didn’t leave.
- *John did leave.

I want John to do leave

- I want John to do not leave

- Do is defective, as are modals, lacking non-finite forms.

- *John does be foolish
- *Does John be foolish
- *John doesn’t be foolish

(48) Be (and have) are not visible in LF, so cannot move to do.