Transderivational Constraints in the Minimalist Program

Gereon Müller

Institut für Linguistik
Universität Leipzig

SoSe 2008

www.uni-leipzig.de/∼muellerg

Fewest Steps: V-to-T Movement in Chomsky (1991)

Background:
Chomsky (1991) is concerned with deriving the difference between French and English with respect to V-to-T movement (cf. Pollock (1989)): French has overt V-to-T movement of finite verbs; English does not have such movement (except for auxiliaries). This is shown in (1).

(1) *V-to-T in French vs. English*:

a. Jean embrasse_1 souvent $[_{\mathrm{VP}}$ t_1 Marie]
b. *Jean souvent $[_{\mathrm{VP}}$ embrasse_1 Marie]
c. *John kisses_1 often $[_{\mathrm{VP}}$ t_1 Mary]
d. John often $[_{\mathrm{VP}}$ kisses_1 Mary]

Questions:
Why do French verbs have to move to T? Why must English main verbs not move to T? Assumption: French has strong T nodes, English has weak T nodes.

(2) *Strength of T^d*:
Strong T tolerates adjunction of all Vs; weak T tolerates adjunction only of "light" Vs (auxiliaries).

Note:
This excludes (1-c) in English: Overt V-to-T movement violates Strength of T. In contrast, overt V-to-T movement in (1-a) in French does not violate this constraint. Still, something extra needs to be said about (1-b) in French, which vacuously fulfills Strength of T, just like (1-d) in English does. Thus, the question is: Why is overt V-to-T raising obligatory if it is possible? Chomsky's (1991) background assumption is that inflection is base-generated in T. If V does not raise to T, T must lower to V in overt syntax, so as to fulfill another local constraint, the *Stray Affix Filter*, which prohibits inflectional affixes that are not attached to a verbal host. The crucial idea now is that overt T lowering creates an unbound trace that must be undone by LF, via covert raising of the whole V-T complex to the position of the trace of T. The derivations underlying (1-a) and (1-b) in French are given in (3) and (4), respectively.

(3) *The V-to-T Derivation in French*:

a. Jean T_2 souvent $[_{VP}$ embrasse$_1$ Marie $]$ (overt raising)
b. Jean $[_{T_2}$ embrasse$_1$-T $]$ souvent $[_{VP}$ t_1 Marie $]$

(4) *The V-in situ Derivation in French*:

a. Jean T_2 souvent $[_{VP}$ embrasse$_1$ Marie $]$ (overt lowering)
b. Jean t_2 souvent $[_{VP}$ $[_{V_1}$ embrasse$_1$-T_2 $]$ Marie $]$ (covert raising)
c. Jean $[_{V_1}$ embrasse$_1$-T_2 $]$ souvent $[_{VP}$ t_1 Marie $]$

Idea:
The second derivation has more movement steps than the first one, and it is therefore filtered out as uneconomical by the transderivational economy constraint *Fewest Steps* that can be formulated as follows:

(5) *Fewest Steps*^{td}:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer operations than D_2, then D_1 is to be preferred over D_2.

(6) *Reference Set*:
Two derivations D_1 and D_2 are in the same reference set iff:

a. D_1 and D_2 start with the same LA.
b. D_1 and D_2 do not violate local or global constraints.

Consequences:
(i) (6-a): (7-b) cannot accidentally block (7-a) even though it involves fewer syntactic operations.
(ii) (6-b): (7-c) cannot accidentally block (7-a) even though it involves fewer syntactic operations by leaving the *wh*-phrase and the auxiliary in situ – (7-c) violates the constraint that selectional features ([*Q*] on C) have to be deleted.

(7) *Motivation for reference sets*:

a. What$_1$ have$_2$ you t$_2$ seen t$_1$?
b. You have seen a car
c. *You have$_2$ seen what$_1$?

Historical comment:
This particular application of a transderivational constraint in the minimalist program is not generally accepted anymore. Still, it can serve as an illustration of certain recurring properties and problems of transderivational constraints, and of competition-based approaches in general.

Problems 1

Problem no. 1:

Chomsky (1991, 433) observes: This system "tends to eliminate the possibility of optionality in derivation. Choice points will be allowable only if the resulting derivations are all minimal in cost ... This may well be too strong a conclusion, raising a problem for the entire approach."

(8) *Optional topicalization*:

a. Mary gave a book to John_1

b. To John_1 Mary gave a book t_1

Solutions 1

Solutions:
(i) Chomsky (1991): Certain movement operations might be "assigned to some other component of the language system, perhaps a 'stylistic' component of the mapping ... to PF;" movement operations of this type might then be exempt from the Fewest Steps constraint.
(ii) We might revise the definition of reference set appropriately, such that the two derivations in (8) do not compete anymore. For instance, we might add the requirement that competing derivations must have identical LF representations, as in (9). Assuming that sentences which differ only with respect to whether topicalization has applied must have different LFs, this would yield the desired result.

(9) *Reference Set* (revised):
Two derivations D_1 and D_2 are in the same reference set iff:

a. D_1 and D_2 start with the same LA and have the same LF representation.
b. D_1 and D_2 do not violate local or global constraints.

Problem no. 2:
Transderivational economy constraints increase complexity: To find out whether a given sentence is grammatical, it does not suffice to look at internal properties of the sentence (Does it violate a local constraint?); rather, the properties of other sentences have to be taken into account as well (Does the sentence have the most economical derivation in the reference set?). Chomsky (1991, 448): “Language design as such appears to be in many respects ‘dysfunctional,’ yielding properties that are not well adapted to the functions language is called upon to perform.”

Problem no. 3:
Successive-cyclic movement: It is standardly assumed that long-distance *wh*-movement must be successive-cyclic; otherwise, a locality constraint will be violated:

(10) *Wh-Island Effects*:
*How$_1$ do you wonder [$_{\text{CP}}$ which car$_2$ to fix t$_2$ t$_1$] ?

(11) *Successive-cyclic movement and Fewest Steps*:

a. How$_1$ do you think [$_{\text{CP}}$ t$_1''$ that John said [$_{\text{CP}}$ t$_1'$ that Bill fixed the car t$_1$]] ?

b. *How$_1$ do you think [$_{\text{CP}}$ that John said [$_{\text{CP}}$ that Bill fixed the car t$_1$]] ?

A caveat:
(10) violates the Subjacency Condition; but given that this constraint is not widely adopted anymore, it is by no means clear how (11-b) can in fact be excluded. As noted before, the main locality constraints adopted in the minimalist program are (a) the Minimal Link Condition (= F-over-F Principle & Superiority Condition), and (b) the Condition on Extraction Domain. The Minimal Link Condition is in fact violated in (10), but not in (11-b); and neither is the Condition on Extraction Domain. How, then, can we ensure that locality constraints demand successive-cyclic movement?

Chomsky's (2000; 2001; 2005; 2006) proposal:
Successive-cyclic movement is forced by the Phase Impenetrability Condition.

(12) *Phase Impenetrability Condition*[d]:
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and SpecX are accessible to such operations.

Note:
The constraint relies on the existence of special derivational units: phases.

(13) *Phase*:
The propositional categories CP and VP/vP are phases; other XPs (except perhaps for DP) are not.

Note:
It follows that movement must be successive-cyclic in an even stricter sense than assumed so far:
(i) Movement from CP must take place via SpecC.
(ii) Movement from VP/vP must take place via SpecV/v.
However, (ii) will be disregarded in what follows. (I.e., for present purposes, we can assume that only CPs are phases.)

A consequence:
Successive-cyclic movement creates a potential problem for the Fewest Steps condition: Successive-cyclic movement as in (11-a) should always be blocked by one-step movement as in (11-b). We would then incorrectly expect all long-distance dependencies to be ill formed.

But:
There is no problem under the notion of reference set in (6) or (9): Only those derivations can compete that respect all local constraints of grammar, i.e., that are otherwise well formed. The derivation that generates the surface representation (11-b) violates the Phase Impenetrability Condition; hence, it cannot compete with the derivation that generates (11-a), and (11-a) is chosen by Fewest Steps because there is no competing derivation that would be more economical.

However:
Suppose that we were to dispense with clause (b) in the definition of reference set, or that clause (b) would be weakened in such a way that some derivations violating local constraints can compete after all. (As we will see below, there is some evidence for this latter option.) If so, the derivations generating (13-a) and (13-b) might compete, and the problem of accounting for successive-cyclic movement under Fewest Steps persists. How, then, can we permit successive-cyclic movement in (13)? We have seen Chomsky's solution already in [3]: Successive-cyclic movement is reinterpreted as a single (albeit complex) operation:

(14) *Form Chain*:
Move α to its target position and freely insert intermediate traces in appropriate positions.

Note:
The French case of T-lowering at S-structure followed by V-raising at LF in (4) would still have to involve two applications of Form Chain.

Wh-Topicalization in Epstein (1992)

Goal:
Epstein accounts for the ban on *wh*-topicalization in English in terms of the transderivational Fewest Steps constraint.

(15) *Optional topicalization: different LF representations*:

a. Who$_1$ t$_1$ said [$_{\text{CP}}$ that [$_{\text{TP}}$ Mary gave a book to John$_2$]] ?
b. Who$_1$ t$_1$ said [$_{\text{CP}}$ that to John$_2$ [$_{\text{TP}}$ Mary gave a book t$_2$]] ?

(16) *The ban on wh-topicalization*:

a. Who$_1$ t$_1$ said [$_{\text{CP}}$ that [$_{\text{TP}}$ Mary gave a book to whom$_2$]] ?
b. *Who$_1$ t$_1$ said [$_{\text{CP}}$ that to whom$_2$ [$_{\text{TP}}$ Mary gave a book t$_2$]] ?

Assumption:
All *wh*-phrases must be in SpecC$_{[+wh]}$ at LF.

(17) *LF representations of (16-a), (16-b)*:
*Who$_1$ to whom$_2$ t$_1$ said [$_{\text{CP}}$ that [$_{\text{TP}}$ Mary gave a book t$_2$]] ?

Analysis

D_1 reaches this LF by applying one (covert) instance of *wh*-movement to the embedded object DP *who*$_2$. There is only one movement operation in this case, either because LF movement of arguments does not have to be successive cyclic, or because successive-cyclic covert movement can be analyzed as one instance of Form Chain. D_2, on the other hand, reaches the same LF by applying two instances of *wh*-movement to the embedded object DP *who*$_2$ (viz., one overtly and one covertly – given the intervening S-structure from which the PF derivation branches, these two movement operations cannot be reanalyzed as one instance of Form Chain). Hence, D_1 blocks D_2 via Fewest Steps.

Note:
The same analysis can be given for ill-formed *wh*-scrambling in German (see Müller & Sternefeld (1996)).

(18) *The ban on wh-scrambling in German*:

a. Warum$_1$ hat der Fritz was$_1$ gelesen ?
why has ART Fritz what read

b. *Warum$_1$ hat was$_1$ der Fritz t$_1$ gelesen ?
why has what ART Fritz read

(19) *The ban on wh-topicalization in German*:

a. Wer$_1$ sagte t$_1$ [$_{\mathrm{CP}}$ dass Maria wem$_2$ ein Buch gegeben hat$_3$] ?
who said that Maria whom a book given has

b. *Wer$_1$ sagte t$_1$ [$_{\mathrm{CP}}$ wem$_2$ hat$_3$ Maria t$_2$ ein Buch gegeben t$_3$] ?
who said whom has Maria a book given

Problem:
German topicalization always requires verb-second movement, which is incompatible with the presence of a complementizer in German. Hence, D_1 generating (19-b) and D_2 generating (19-a) do not share an identical lexical array. Solution:

(20) *Reference Set* (second revision):
Two derivations D_1 and D_2 are in the same reference set iff:

a. D_1 and D_2 have the same LF representation.
b. D_1 and D_2 do not violate local or global constraints.

Assumption:
A complementizer *dass* is deleted (or invisible) at LF.

Another problem:
Optional partial *wh*-movement to a $\text{SpecC}_{[-wh]}$ position in Ancash Quetchua (see Cole (1982)).

(21) *Partial wh-movement in Ancash Quechua*:

a. $[_{\text{CP}}$ Ima-ta-taq$_1$ (qam) kreinki $[_{\text{CP}}$ t''_2 Maria muna-nqa-n-ta $[_{\text{CP}}$ t'_1
what$_{acc}$ you believe Maria want-nom-3-acc
José t_1 ranti-na-n-ta]]] ?
José buy-nom-3-acc

b. $[_{\text{CP}}$ – (Qam) kreinki $[_{\text{CP}}$ ima-ta-ta$_1$ Maria muna-nqa-n-ta $[_{\text{CP}}$ t'_1 José t_1 ranti-na-n-ta]]] ?

c. $[_{\text{CP}}$ – (Qam) kreinki $[_{\text{CP}}$ – Maria muna-nqa-n-ta $[_{\text{CP}}$ ima-ta-ta$_1$ José t_1 ranti-na-n-ta]]] ?

d. $[_{\text{CP}}$ – (Qam) kreinki $[_{\text{CP}}$ – Maria muna-nqa-n-ta $[_{\text{CP}}$ –José ima-ta-ta$_1$ ranti-na-n-ta]]] ?

(22) *Reference Set* (third revision):
Two derivations D_1 and D_2 are in the same reference set iff:

a. D_1 and D_2 have the same S-structure representation.
b. D_1 and D_2 do not violate local or global constraints.

Result:
Partial *wh*-movement is permitted, but most of the original evidence in favour of Fewest Steps is lost: Thus, on this view, neither French V-in situ, nor English (or German) *wh*-topicalization can be ruled out by Fewest Steps anymore. As noted in Sternefeld (1997), this situation might be viewed as indicative of a general problem with transderivational constraints: A significant reduction of competition in reference sets may be empirically desirable so as to account for cases of optionality (as in partial *wh*-movement constructions); but as an unwanted side effect, it also threatens to undermine the notion of transderivational economy itself: Many ill-formed derivations that could be ruled out by transderivational constraints will now survive because the more economical derivation is not part of the same reference set anymore. Finding a suitable definition of reference set that is weak enough to permit optionality and strong enough to actually do some work is one of the fundamental concerns of all versions of the minimalist program that employ the notion of competition.

Freezing in Collins (1994)

Background:
Evidence for yet another definition of reference sets comes from Collins' (1994) account of freezing effects with A-movement in English. We have seen that subject DPs are islands in English, whereas complement DPs permit extraction (with certain types of verbs). An interesting case is that of subject DPs that originate in complement position, as in the case of passivization. As can be seen in (23-c), such derived subject DPs are also islands.

(23) *Derived subject DPs as islands*:

a. Who$_1$ did John take [$_{\mathrm{DP}}$ a picture of t$_1$] ?
b. *Who$_1$ is [$_{\mathrm{DP}}$ a picture of t$_1$] on sale ?
c. *Who$_1$ was [$_{\mathrm{DP}_2}$ a picture of t$_1$] taken t$_2$ by John ?

(24) *D_1 violates the Condition on Extraction Domain*:

a. $[_{\mathrm{VP}}$ taken $[_{\mathrm{DP}_2}$ a picture of who$_1$] by John] ...
b. $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of who$_1$] T $[_{\mathrm{VP}}$ taken t$_2$ by John]]
c. *$[_{\mathrm{CP}}$ who$_1$ was $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of t$_1$] T $[_{\mathrm{VP}}$ taken t$_2$ by John]]]

(25) *Condition on Extraction Domain*:

a. Movement must not cross a barrier.
b. An XP is a barrier iff it is not a complement.

(26) *D_2 violates the Strict Cycle Condition*:

a. $[_{\text{VP}}$ taken $[_{\text{DP}_2}$ a picture of who$_1$] by John] ...

b. $[_{\text{CP}}$ who$_1$ was $[_{\text{TP}}$ T $[_{\text{VP}}$ taken $[_{\text{DP}_2}$ a picture of t$_1$] by John]]]

c. $[_{\text{CP}}$ who$_1$ was $[_{\text{TP}}$ $[_{\text{DP}_2}$ a picture of t$_1$] T $[_{\text{VP}}$ taken t$_2$ by John]]]

(27) *Strict Cycle Condition*[d] (simplified):
No operation can apply to a domain dominated by a cyclic node α in such a way as to affect solely a proper subdomain of α dominated by a node β which is also a cyclic node.

(28) *D_3 exhibits chain interleaving*:

a. $[_{\mathrm{VP}}$ taken $[_{\mathrm{DP}_2}$ a picture of who_1] by John]
b. $[_{\mathrm{VP}}$ who_1 $[_{\mathrm{V'}}$ taken $[_{\mathrm{DP}_2}$ a picture of t_1] by John]]
c. $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of t_1] T $[_{\mathrm{VP}}$ who_1 $[_{\mathrm{V'}}$ taken t_2 by John]]]
d. $[_{\mathrm{CP}}$ who_1 was $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of t_1] T $[_{\mathrm{VP}}$ t'_1 $[_{\mathrm{V'}}$ taken t_2 by John]]]]

Problem:
How is chain interleaving in D_3 ruled out?
Solution:
D_3 is blocked by D_1 and D_2 via Fewest Steps: Other things being equal, D_3 needs three movement steps where D_1 and D_2 make do with two movement steps.

Note:
The argument presupposes that the first movement operation in D_3 can somehow respect the Economy Constraint on Merge (i.e., there must be a selectional feature that is deleted by the Move operation to SpecV).

Consequence:
We need yet another definition of reference sets. The three derivations D_1, D_2, and D_3 yield the same surface string, which is ill formed. Thus, the more economical derivations that D_3 is blocked by via Fewest Steps are not well-formed derivations, as in the applications of Fewest Steps discussed above, but rather ill-formed derivations that violate local constraints. This reasoning implies that reference sets can in fact not be defined as assumed so far, by requiring that only those derivations can compete that satisfy all local constraints – in the case at hand, D_1 and D_2 violate local constraints. Still, we cannot simply drop this requirement in the definition of reference sets; otherwise, all instances of movement would invariably be blocked in favour of in-situ derivations by Fewest Steps, and syntactic derivations would be fairly trivial. It seems that what is needed in view of this conflicting evidence is a relativized notion of local constraint satisfaction.

The solution: convergence (Chomsky (1993; 1995)):
Only those derivations can compete with respect to transderivational constraints that converge. Essentially, whereas all violations of local constraints lead to ungrammaticality, only a subset of violations of local constraints also lead to non-convergence. Ungrammatical derivations that converge may then still be used to block other derivations as ungrammatical, as in the freezing construction discussed by Collins (1994).

(29) *Convergence*:

a. Violations of constraints related to selectional features and their deletion lead to non-convergence.
b. Violations of locality constraints and the Strict Cycle Condition lead to convergence.

(30) *Reference Set* (fourth revision):
Two derivations D_1 and D_2 are in the same reference set iff:

a. D_1 and D_2 have the same LA/S-structure/LF.
b. D_1 and D_2 converge.

Shortest Paths

(31) *Shortest Paths*^{td} (Chomsky (1993; 1995)):
If two derivations D_1 and D_2 are in the same reference set and the movement paths of D_1 are shorter than the movement paths of D_2, then D_1 is to be preferred over D_2.

Superiority Effects in Chomsky (1993) and Kitahara (1993)

(32) *Superiority effects revisited*:

a. I wonder [$_{\mathrm{CP}}$ who$_1$ C [$_{\mathrm{TP}}$ t$_1$ bought what$_2$]]
b. *I wonder [$_{\mathrm{CP}}$ what$_2$ C [$_{\mathrm{TP}}$ who$_1$ bought t$_2$]]
c. Whom$_1$ did John persuade t$_1$ [$_{\mathrm{CP}}$ to visit whom$_2$] ?
d. *Whom$_2$ did John persuade whom$_1$ [$_{\mathrm{CP}}$ t$'_2$ to visit t$_2$] ?

(33) *Superiority Condition*[d] (Chomsky (1973)):
In a structure $\alpha_{[*F*]}$... [... $\beta_{[F]}$ ... [... $\gamma_{[F]}$ ...] ...] ..., movement to [*F*] can only affect the category bearing the [F] feature that is closer to [*F*].

Assumption:
Superiority effects can be accounted for by the transderivational condition Shortest Paths. For instance, the movement path from t$_1$ to *whom*$_1$ in (32-c) is shorter than the movement path from t$_2$ to *whom*$_2$ in (32-d).

Problem:
At least some of the evidence for Fewest Steps (the ban on V-in situ in French, the ban on *wh*-topicalization and *wh*-scrambling in English and German) has relied on the assumption that covert movement counts in the same way that overt movement does. But assuming that LF movement also counts for the Shortest Paths condition leads straightforwardly into a dilemma: In the case at hand, the derivation that has the shorter overt *wh*-movement path invariably has the longer covert *wh*-movement path, and it seems that by LF, both derivations have *wh*-movement paths of equal length. Hence, ceteris paribus, both should be well formed.

Solutions:
(i) There is in fact no covert *wh*-movement of any kind (this is incompatible with Fewest Steps approach outline above).
(ii) Whereas Fewest Steps compares whole derivations, Shortest Paths compares only the overt parts of derivations.

Another Problem

LF-optionality: Sentences like (34-a) have two possible readings (see Baker (1970)) that correspond to two different LF representations, given LF movement of *wh*-in situ elements.

(34) *Baker sentences*: Who$_1$ t$_1$ wonders [$_{\mathrm{CP}}$ where$_2$ we bought what$_3$ t$_2$] ?

a. who$_1$ what$_3$ t$_1$ wonders [$_{\mathrm{CP}}$ where$_2$ we bought t$_3$ t$_2$]
Answer: John wonders where we bought the books, Mary wonders where we bought the records, etc.

b. who$_1$ t$_1$ wonders [$_{\mathrm{CP}}$ where$_2$ what$_3$ we bought t$_3$ t$_2$]
Answer: John wonders where we bought what, Mary wonders where we bought what, etc.

Problem:
Given that all *wh*-in situ phrases must undergo movement to a SpecC$_{[+wh]}$ position at LF, D$_2$ (creating (34-b)) should block D$_1$ (creating (34-a)) because D$_1$'s paths are longer.

Solutions:
(i) As before: Covert *wh*-movement either does not exist, or does not count with respect to the Shortest Paths condition.
(ii) Reference sets are defined in such a way that competing derivations must have identical LF representations.

Yo-Yo Movement in Collins (1994)

Note:
Yo-yo movement characterizes a combination of lowering and raising operations affecting a single item in the course of a derivation. Collins (1994) shows that the availability of yo-yo movement would make a wrong prediction for the West African language Ewe, and attempts to derive a ban on yo-yo movement from the Shortest Paths condition.

Background

Ewe is among the languages that show reflexes of of successive-cyclic *wh*-movement in the C domain. The reflex of successive cyclicity concerns the morphological form of the 3.Pers.Sing. subject pronoun in the canonical subject position. The regular form of the pronoun is *é*; cf. (35-a). The regular pronoun *é* can be replaced by *wo* in cases of long-distance extraction (focus movement, in the case at hand); cf. (35-b).

(35) *Successive cyclicity in Ewe*:

a. Kofi gblǒ [$_{\text{CP}}$ be é/*wo ʃo Kǒsi]
   Kofi said that he hit Kǒsi

b. Kofi$_1$ ϵ me gblǒ [$_{\text{CP}}$ (t$'_1$) be é/wo ʃo t$_1$]
   Kofi Foc I said that he hit

Assumptions:
(i) *é* is replaced by *wo* iff the local SpecC position is filled.
(ii) Long-distance movement of argument DPs in Ewe may apply either successive-cyclically, via SpecC (in which case *wo* is obligatory), or in one step (in which case *wo* is impossible).

(36) *Movement originating in the matrix CP*:

a. Kofi$_1$ ϵ me gblŏ na t$_1$ [$_{\mathrm{CP}}$ be é/*wo ʃo Kŏsi]
Kofi Foc I said to that he hit Kŏsi

b. *A derivation that involves lowering*:

(i) [$_{\mathrm{VP}}$ said [$_{\mathrm{PP}}$ to Kofi$_1$] [$_{\mathrm{CP}}$ that [$_{\mathrm{TP}}$ he hit Kŏsi]]]

(ii) [$_{\mathrm{VP}}$ said [$_{\mathrm{PP}}$ to t$_1$] [$_{\mathrm{CP}}$ Kofi$_1$ that [$_{\mathrm{TP}}$ he hit Kŏsi]]] ...

(iii) [$_{\mathrm{CP}}$ Kofi$_1$ Foc [$_{\mathrm{TP}}$ T [$_{\mathrm{VP}}$ said [$_{\mathrm{PP}}$ to t$_1$] [$_{\mathrm{CP}}$ t_1' that [$_{\mathrm{TP}}$ he hit Kŏsi]]]]]

Note:
The illformedness of this derivation also follows from the Strict Cycle Condition. Still, Collins (1994) observes that D_1 in (36) is blocked by the derivation D_2 in (37) via Shortest Paths. D_2 proceeds without intermediate lowering.
(One might think that Fewest Steps would also suffice to block D_1 in favour of D_2. However, assuming that the two movement operations in D_1 can be reanalyzed as a single instance of Form Chain, this is not the case.)

(37) *A derivation that does not involve lowering*:

a. $[_{\mathrm{VP}}$ said $[_{\mathrm{PP}}$ to Kofi$_1$] $[_{\mathrm{CP}}$ that $[_{\mathrm{TP}}$ he hit Kǒsi]]]

b. $[_{\mathrm{CP}}$ Kofi$_1$ Foc $[_{\mathrm{TP}}$ T $[_{\mathrm{VP}}$ said $[_{\mathrm{PP}}$ to t$_1$] $[_{\mathrm{CP}}$ that $[_{\mathrm{TP}}$ he hit Kǒsi]]]]]

A consequence for the definition of reference set:
"He" is a *wo* in D_1 and an *é* in D_2. Hence, this difference must not suffice to create different reference sets.

Tagalog Wh-Movement in Nakamura (1998)

Generalization:
Only the highest A-position of a given clause (the subject position SpecT) is accessible for *wh*-movement in the Austronesian language Tagalog. In constructions in which an agent DP occupies the highest A-position (the so-called *Agent Topic* (AT) constructions), this DP can be moved; an DP bearing a different Theta-role that shows up in an object position cannot undergo *wh*-movement.

(38) *Wh-movment in AT constructions*:

a. $[_{\text{CP}}$ Sino$_1$ ang $[_{\text{TP}}$ t'_1 b-um-ili $[_{\text{VP}}$ t_1 t_V ng damit$_2$]]] ?
who Ang bought$_{AT}$ dress$_{inh}$
'Who is the one that bought the dress?'

b. *$[_{\text{CP}}$ Ano$_2$ ang $[_{\text{TP}}$ *si Juan*$_1$ b-um-ili $[_{\text{VP}}$ t_1 t_V t_2]]]$_?$
what Ang Juan$_{abs}$ bought$_{AT}$
'What is the thing that Juan bought?'

Observation:
A different marking on the verb triggers the so-called *Theme Topic* (TT) construction. Here, the theme DP occupies the structural subject position SpecT; and indeed, only the theme DP can undergo *wh*-movement.

(39) *Wh-movment in TT constructions*:

a. *[$_{\text{CP}}$ Sino$_1$ ang [$_{\text{TP}}$ *ang damit*$_2$ b-in-ili [$_{\text{VP}}$ t_1 t_V t_2]]] ?
who Ang dress$_{abs}$ bought$_{TT}$
'Who is the one that bought the dress?'

b. [$_{\text{CP}}$ Ano$_2$ ang [$_{\text{TP}}$ t'_2 b-in-ili [$_{\text{VP}}$ ni Juan t_V t_2]]] ?
what Ang bought$_{TT}$ Juan$_{erg}$
'What is the thing that Juan bought?'

Idea:
The derivations generating (38-a) and (39-a) compete, as do the derivations generating (38-b) and (39-b). The derivations underlying (38-a) and (39-b) can then block their respective competitors as ungrammatical because of the Shortest Paths constraint. Consider *wh*-movement of the theme DP in (38-b) and (39-b). The movement path from the VP-internal object position to the SpecC target position in (38-b) is longer than the path from the subject position SpecT to SpecC in (39-b).

Consequences:
(i) Derivations can compete even though they do not have identical lexical material – the Agent Topic and the Theme Topic constructions clearly differ in lexical make-up: "identical lexical array" is replaced by "non-distinct lexical array" in the definition of reference sets. The latter is defined in such a way that two lexical arrays that only differ with respect to functional features do not count as distinct.
(ii) The derivation that generates (39-b) may minimize the *wh*-path in comparison with the derivation that generates (38-b), but it increases path lengths in the TP-domain. Nakamura (1998) replaces the notion of "movement paths" in the definition of the Shortest Paths condition with the more specific notion of "comparable chain links".

Freezing in Chomsky (1995)

(40) *Derived subject DP islands again*:
*Who$_1$ was [$_{\mathrm{DP}_2}$ a picture of t$_1$] taken t$_2$ by John ?

Recall from the discussion of Collins (1994):
(i) D_1 violates the Condition on Extraction Domain.
(ii) D_2 violates the Strict Cycle Condition.
(iii) D_3 violates the Fewest Steps condition.

Proposal:
D_2 can be accounted for without invoking the Strict Cycle Condition, by invoking the Shortest Paths Condition.

(41) a. *D_1: Condition on Extraction Domain*:

(i) $[_{\mathrm{VP}}$ taken $[_{\mathrm{DP}_2}$ a picture of who_1] by John] ...

(ii) $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of who_1] T $[_{\mathrm{VP}}$ taken t_2 by John]]

(iii) *$[_{\mathrm{CP}}$ who_1 was $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of t_1] T $[_{\mathrm{VP}}$ taken t_2 by John]]]

b. *D_2: Shortest Paths Condition*:

(i) $[_{\mathrm{VP}}$ taken $[_{\mathrm{DP}_2}$ a picture of who_1] by John]] ...

(ii) $[_{\mathrm{CP}}$ who_1 was $[_{\mathrm{TP}}$ T $[_{\mathrm{VP}}$ taken $[_{\mathrm{DP}_2}$ a picture of t_1] by John]]]

(iii) $[_{\mathrm{CP}}$ who_1 was $[_{\mathrm{TP}}$ $[_{\mathrm{DP}_2}$ a picture of t_1] T $[_{\mathrm{VP}}$ taken t_2 by John]]

Analysis (Chomskys (1995)):
"Passive [i.e., DP raising] is the same in both [derivations]; *wh*-movement is 'longer' in the illicit one in an obvious sense, object being more remote from SpecC than subject in terms of number of XPs crossed. The distinction might be captured by a proper theory of economy of derivation." D_2 in (41-b), which does not violate a local constraint (the SCC, by assumption, being irrelevant), is blocked as ungrammatical via Shortest Paths by the more economical D_1 in (41-a), which does violate a local constraint (the CED) but converges.

Merge before Move

Note:
Given that Merge and Move alternate in the derivation, the situation can arise that it must be decided whether the next step is a Merge or a Move operation. The following transderivational condition settles the issue by preferring Merge to Move if both are possible as such; the specific formulation is based on Frampton & Gutman (1999).

(42) *Merge before Move^{td}*:
Suppose that two derivations D_1 and D_2 are in the same reference set and respect all local constraints, and $D_1 = <\Sigma_0, ..., \Sigma_n, \Sigma_{n+1}, ...\Sigma_k>$ and $D_2 = <\Sigma_0, ..., \Sigma_n, \Sigma'_{n+1}, ...\Sigma'_{k'}>$. Then D_1 is to be preferred over D_2 if $\Sigma_n \rightarrow \Sigma_{n+1}$ is an instance of Merge and $\Sigma_n \rightarrow \Sigma'_{n+1}$ is an instance of Move.

(43) *Expletive constructions in English*:

a. There$_1$ seems [$_{\text{TP}}$ t$_1$ to be [$_{\text{PP}}$ someone$_2$ in the room]]

b. *There$_1$ seems [$_{\text{TP}}$ someone$_2$ to be [$_{\text{PP}}$ t$_2$ in the room]]

Assumption:
Non-finite T in a control infinitive has a [*D*] feature (as does the finite T in the matrix clause).

Note:
The lexical array is identical in (43-a) and (43-b). The transderivational constraint Merge before Move ensures that D_1 generating (43-a) is to be preferred over D_2 generating (43-b).

Question:
Why is (44) possible, which does not have expletive Merge at all?

(44) *Successive-cyclic movement to SpecT*:
Someone$_1$ seems [$_{\mathrm{TP}}$ t$'_1$ to be t$_1$ in the room]

Answer:
Reference sets must be assumed to be defined via identical lexical arrays. The derivation that generates (44) has a lexical array that is different from the one underlying the derivations in (43).

Conceptual problem:
Why should it be that Merge is preferred over Move?

Solution:
Move is defined on the basis of Move, which makes it inherently more complex. Chomsky (2000): “Good design conditions would lead us to expect that simpler operations are preferred to more complex ones, so that Merge ... preempt[s] Move, which is a ‘last resort,’ chosen when nothing else is possible.”

Conclusion

Note:
There are more transderivational (or translocal) constraints that have been suggested in the minimalist literature. Among them we find:

1. Procrastinate (transderivational, Chomsky (1993; 1995));
2. Economy of Representation (translocal, Chomsky (1991));
3. Preference Principle for Reconstruction (transderivational, Chomsky (1993)).

A Meta-Theoretical Consideration

Grammars that employ transderivational (or translocal) constraints in addition to local (or local and global) constraints all have the format that Prince & Smolensky (1993) attribute to an optimality-theoretic grammar: A first generator part (called *Gen*) creates the *candidate set* (= reference set, in minimalist syntax); Gen has only local (or global) constraints. A second harmony-evaluation part (called *H-Eval*) determines the optimal candidate(s) (= derivation(s), in minimalist syntax) in the candidate set. More generally, all kinds of competition-based syntax have this structure, which is schematically shown in (45).

(45) *Structure of a competition-based syntax*:

a. Gen creates the candidate set $\{C_1, C_2, ...\}$.

b. H-Eval determines the optimal candidate(s) C_i (C_j, ...) in $\{C_1, C_2, ...\}$.

Question:
Does optimality equal grammaticality in the minimalist program?

Question:
Does optimality equal grammaticality in the minimalist program?

Answer:
Not necessarily: Given the notion of convergence, an optimal candidate may be one that violates a local (or global) constraint, and is therefore ungrammatical.

Problem:
How does optimality evaluation proceed in the presence of more than one transderivational constraint in the grammar? In this case, conflicts may arise. As a simple, abstract example, suppose that there are two transderivational constraints (T_1 and T_2), and only three derivations (D_1, D_2, and D_3) in the reference set. Suppose further that T_1 prefers D_1 over D_2 and D_3; that T_2 prefers D_2 over D_1 and D_3; and that a derivation D_0 that would be preferred by both T_1 and T_2 fails to converge, so that it cannot participate in the competition. In such a situation, various possibilities arise:

Conflicting Transderivational Constraints

1. *Tolerance*:
   It suffices to be selected by one transderivational constraint to be optimal; consequently, both D_1 and D_2 would be classified as optimal.

Conflicting Transderivational Constraints

1. *Tolerance*:
   It suffices to be selected by one transderivational constraint to be optimal; consequently, both D_1 and D_2 would be classified as optimal.
2. *Ranking*:
   The conflict among transderivational constraints is resolved by a ranking, such that that derivation is optimal that is preferred by the higher-ranked constraint in the case of conflict. If T_1 is ranked higher than T_2, this would imply that only D_1 is optimal.
3. *Breakdown*:
   In the case of conflicting instructions made by transderivational constraints, no derivation can emerge as optimal.

Note:
The last option is generally adopted in the literature (see Collins (1994), Sternefeld (1997), and Müller (2000)).

(46) *Grammaticality in minimalist syntax*:
A derivation D_i is grammatical iff (a) and (b) hold:

a. D_i does not violate a local constraint.
b. D_i is optimal.

(47) *Optimality in minimalist syntax*:
A derivation D_i is optimal iff there is no derivation D_j in the same reference set that is preferred over D_i by a transderivational or translocal constraint.

Overall Conclusion

Transderivational constraints can offer insightful analyses, but they are problematic from the point of view of complexity (competition in potentially infinite candidate/reference sets). Hence: There is a strong tendency in recent versions of the minimalist program to dispense with transderivational constraints – and hence, with the concept of competition – altogether; see in particular Collins (1997), Frampton & Gutman (1999), and also Chomsky (1995; 2000; 2001). However, the transderivational Merge before Move constraint is still widely adopted (see, e.g., Chomsky (2000; 2001)).

A Historical Note

It is interesting to note that the fall of transderivational constraints (and with it the fall of the concept of competition) in minimalist syntax goes hand in hand with the rise of optimality theory, and hence optimality-theoretic syntax, which inherently relies on transderivationality and competition.