'Shortest Move' vs. 'Fewest Steps'

C. Jan-Wouter Zwart

0. In much recent work in generative grammar, derivations are considered to be subject to principles of economy. In Chomsky (1992), economy of derivation is implemented in at least two ways: derivations should involve the shortest possible movements and the fewest possible steps. As Chomsky notes, these two requirements appear to be contradictory. He proposes a *Form Chain* mechanism to resolve the contradiction.

In this paper, I will argue that the Form Chain mechanism consists of two processes, one of which generates empty wh-elements in the intermediate specifier positions required for long distance movement/chain formation. The other process consists in movement *across* these intermediate wh-elements, in compliance with the fewest steps requirement, but in violation of the shortest move requirement. I conclude that the shortest move requirement does not apply in long distance movement processes.

In the remainder of this paper, I will show that the shortest move requirement is equally redundant in the analysis of the other phenomena that seem to call for it, head movement and superraising. My conclusion will be that economy of derivation consists in the single requirement that the number of movement steps should be as small as possible.[1]

1. Let us start by assuming that subjects are generated inside VP, the VP constituting a kind of kernel sentence consisting of a subject, a verb, and an object. Let us also assume that the elements of a kernel sentence have to be formally licensed in well-defined syntactic configurations, called *functional projections*. Syntax then consists in moving the elements of a kernel sentence to designated positions in functional projections. The output of the syntactic component is an interface representation serving as the input to other cognitive systems, such as those involved in speech processing and interpretation.

As a minimalist principle, Chomsky 1992 assumes that the interface representations should be pure and simple, stripped of all features that are not relevant to the cognitive systems they provide input for. This he calls *economy of representation*, summarized in (1):[2]

(1) *Economy of representation*:
Use as few symbols as possible in the output of a derivation

In addition to (1), Chomsky proposes a second minimalist principle, stating that interface representations should be arrived at in the most economical way. This paper

[1] This material was presented at the TABU-dag at the University of Groningen, June 25, 1993, at the Stafseminar of the Department of Linguistics of the University of Tilburg, September 8, 1993, and at the Research Seminar of the Department of Linguistics of the University of Geneva, March 8, 1994. I would like the audiences at these occasions, in particular Jaume Solà, Henk van Riemsdijk, Riny Huybregts, Luigi Rizzi, Ur Shlonsky, and Liliane Haegeman. Also thanks to Eric Hoekstra, Phil Branigan, Gereon Müller, Anthony Kroch, Erich Groat, and K. Scott Ferguson.

[2] Economy of representation ultimately reduces to the principle of *Full Interpretation*, formulated in (28) below.

discusses the proper formulation of this second principle, called *economy of derivation*. I will argue for the following formulation:

(2) *Economy of derivation*:
Use as few steps as possible in deriving an output representation

Not much is new in (2). Chomsky 1992 argues that derivations are governed by principles summarized here under the label *inertness*:

(3) *Procrastinate*:
Move as late as possible

(4) *Greed*:
Move only to contribute to personal licensing

(3) and (4) can be grouped together as in (5):[3]

(5) *Inertness*
Move as little as possible

(5) and (2) are equivalent.

The formulation of economy of derivation in (2) is more interesting for what it leaves out than for what it contains. In particular, (2) makes no reference to the *length* of the steps involved in a derivation. According to conventional wisdom, short steps are more economical than long steps. Thus, we could imagine that economy of derivation contains (6) in addition to (2):

(6) *Economy of derivation part 2*:
In deriving a representation, make the shortest possible movements

(6) underlies the concept of *minimality* (Chomsky 1986, Rizzi 1990), paraphrased in (7):

(7) *Minimality*:
In a derivation, don't move across a place where you could have landed

(6) also plays a major part in Chomsky 1992.

2. To illustrate the workings of (6) in Chomsky 1992, consider the derivation of a simple sentence consisting of a subject, a verb, and an object:

(8) John loves Mary

Let us assume that all elements in (8) have been moved out of their kernel sentence positions to their licensing positions.[4] Let us also assume that the architecture of the functional domain is as in (9), where subjects are licensed in the spec position in

[3] The principle *Procrastinate* reduces to *Inertness*, if the latter is considered to apply at each point in the derivation.

[4] It is irrelevant whether this is the correct analysis of English sentences.

AgrSP, objects in the spec position in AgrOP, and the finite verb in the head position in TP, and where VP equals the kernel sentence:

(9) [AgrSP spec head [TP spec head [AgrOP spec head [VP SUB V OB]]]]

In the syntactic derivation of (8), the object has to move from the position indicated by OB in (9) to the spec position in AgrOP. In doing so, the object crosses the position indicated by SUB in (9). Assuming this to be a position where the object could have landed, movement of OB to the spec position in AgrOP violates the shortest steps requirement (6).

A strict application of (6), then, makes all object movement impossible. To remedy this problem, Chomsky proposes that overt movement of the verb (V in (9)) to the head position of AgrOP makes the position indicated by SUB in (9) and the spec position of AgrOP 'equidistant'. Spec,AgrOP and SUB being equidistant, movement of the object to Spec,AgrOP is in keeping with (6). Verb movement, then, is the way to make the VP transparent.

As Chomsky notes, this analysis makes the prediction that overt object movement never occurs without overt verb movement to the head of AgrOP. But this is clearly wrong for languages like German and Dutch, as can be concluded from Vikner 1991, section 4.2.5. In German and Dutch, objects always leave the VP, which must be described as movement to Spec,AgrOP in a minimalist approach. Since the object and the verb are not adjacent (in embedded clauses), the verb cannot have moved to the head of AgrOP in these constructions:

(10) ..daß Peter das Buch gestern nicht gekauft hat
that Peter the book yesterday not bought has

The only way to align the Germanic object movement facts with Chomsky's equidistance principle is to assume that the head of AgrOP is located to the right of the VP, an option excluded in minimalist work such as Zwart 1992b, 1993a, 1993b, and Kayne 1993.

We must therefore conclude that the equidistance principle is problematic. In what follows, I will argue that the shortest move requirement is not a part of economy of derivation. Since the equidistance principle was prompted by the shortest move requirement, abolishing the latter removes the need for the former, thereby solving the problems it poses.

3. There are several phenomena for which the shortest move requirement appears to be relevant. These include head movement, superraising, and wh-movement. I will discuss these phenomena one by one, arguing that the shortest move requirement is irrelevant for the analysis of their properties.

4. The history of the *Head Movement Constraint* (Travis 1984:131) is instructive for our purpose:

(11) *Head Movement Constraint*:
An X° may only move into the Y° which properly governs it

This constraint bars the derivation of (12a) from (12b), where *kiss* and *will* are both X°s (heads), and one of these heads has to move to C:

(12) a. * Who have John will kissed?
b. Who C [John will [have [kissed]]]
c. Who will John have kissed?

As (12c) shows, only the higher auxiliary *will* may move to C.

From the outset it has been clear that the Head Movement Constraint is part of a principle with a larger scope, the *Empty Category Principle* (see Travis 1984:133, Chomsky 1986):

(13) *Empty Category Principle*:
Empty categories must be properly governed

A trace is *properly governed* if it is governed by its antecedent.[5] In (12b), movement of *kiss* to C would leave a trace which is not properly governed by its antecedent (*kiss*), since the trace would be separated from its antecedent by two maximal projections functioning as *barriers* for government (indicated by the brackets in 12b).[6]

Chomsky 1991 argues that the Head Movement Constraint, if it is reducible to the Empty Category Principle, can be dismissed ''as a descriptive artefact, valid only insofar as it does in fact reduce to the ECP.'' Importantly, Chomsky (1991:429), continuing where Chomsky 1986 (p.88) left off, considers the ECP as a *condition on chains*.

This suggests that in (12) the process of deriving (12a) from (12b) is legitimate, whereas the resulting representation is failing. The phenomena associated with the Head Movement Constraint thus traditionally fall in the domain of conditions on representations.

In Chomsky 1992, government does not play a role, and hence conditions on head movement cannot be reduced to (11) or (13). It appears that here, the shortest move requirement of economy of derivation becomes crucial. It takes a longer step to derive (12a) from (12b) than it does (12c).

However, closer scrutiny learns that (12a) can be excluded in a more minimalist way. The question that has to be asked first, namely, is, What is the trigger for verb movement to C in wh-constructions? If this trigger does not apply to infinitives, (12a) will never be derived because of the inertness principle.

There is ample evidence that verb movement to C in Germanic is closely linked to tense. Consider the following facts from Dutch:

(14) Koopt Jan een huis?
(15) a. Jan een huis kopen?
b. * Kopen Jan een huis?

Assuming that the structure of yes/no questions matches that of wh-questions, (14) and (15) are comparable to (12). We may consider the counterpart to the wh-word in (12) to be empty in (14) and (15). This suggests that the verb movement in (14), as in

[5] I ignore θ-government for expository purposes, cf. Chomsky 1986:69.

[6] This is the analysis in Chomsky 1986, Ch. 11.

(12c), targets C. As can be seen in (15), such verb movement only takes place when the verb is finite.

In terms of Chomsky 1992, we may suppose that C hosts a tense feature, comparable to the V-features of AgrS etc., which must be checked by moving T(ense) to C (cf. Wilder and Ćavar 1993). This triggers movement of the finite auxiliary in (12).

The proper test case for the Head Movement Constraint, then, contains a choice of two finite verbs, only the higher of which may be moved to C:

(16) a. John did not think I could help someone
b. Who did John t not think I could help?
c. * Who could John did not think I t help?

(16c) is correctly excluded by the Head Movement Constraint and the shortest move requirement. However, the ungrammaticality of (16c) also follows from an independently needed locality condition, namely that an element can only escape from a tensed clause via the specifier position of CP (Chomsky 1973, 1986). Head movement cannot proceed through a specifier position by definition (and by requirements of uniformity of chains). (16), then, does not suffice to establish the shortest move requirement.

More generally, the observation that heads move stepwise rather than long distance is explained by the feature checking requirements of the minimalist approach. If a functional head Y contains a feature to be checked by the verb X, and X moves to a functional head Z, crossing Y, the feature in Y will remain unchecked, and the derivation will not converge. This covers the core cases of the Head Movement Constraint. Hence, for deriving stepwise head movement, no special constraint needs to be formulated.[7]

Taking (12) and (16) to be representative of Head Movement Constraint phenomena, this analysis makes the shortest move requirement redundant for the proper understanding of this class of facts.

5. The shortest move requirement likewise appears to play a role in barring *superraising* phenomena:

(17) a. *John seems is likely to win
b. *John seems it is likely to win

The sentences in (17) are derived from more basic representations in which *John* is the subject of *win*, generated inside the VP as previously assumed. As (18) shows, the subject position (Spec,AgrS) of the embedded clause is a legitimate target for subject movement:

(18) It seems John is likely to win

[7] The reduction of the Head Movement Constraint to general requirements of feature checking predicts that if no features in a head Y need to be checked by the verb X, X may move to Z crossing Y. In Zwart (1993b) I discuss several of these cases (see also Ouhalla 1989).

It seems, then, that the sentences in (17) are derived by moving *John* across a legitimate target for subject movement, in violation of the shortest move requirement of economy of derivation (Chomsky 1992:21).

However, it is immediately obvious that (17a), at least, is excluded on standard minimalist assumptions of movement and feature checking. Chomsky 1992 argues that all movement operations are triggered by feature checking requirements, and that a given functional head contains two features: one for checking against the features of an XP (*N-feature*) and another one for checking against the features of an X° (*V-feature*). These features are only relevant for syntax, and therefore have to be eliminated before the derivation reaches the interface levels. Movement to a head position or to a specifier position in the functional domain creates the configurations in which matching features can be checked (cf. Zwart 1992). Features, once checked, are automatically eliminated; unchecked features yield a 'crashing' derivation.[8]

Let us assume that in (18), the subject, *John*, moves to the specifier position of AgrS. This creates the proper configuration for checking the N-features of AgrS against the corresponding features of the subject.[9] As a result, the N-features of AgrS and the corresponding features on the subject are eliminated.

Consider the consequences for (17a). Suppose (17a) is derived from an intermediate representation corresponding to (18):

(17a') seems John is likely to win

At this point, the subject's features will be eliminated as a result of the checking operation taking place in the AgrSP in the embedded clause, just like in (18). This will make it impossible for (17a) to be derived from (17a'): the subject has lost its features, and cannot serve to check the features of the AgrS in the matrix clause. This derivation of (17a) is thus excluded.

Another possible derivation of (17a) moves the subject in one swoop from the VP-internal position to the specifier position of the matrix AgrSP. This creates the proper configuration for checking the N-features of the matrix AgrSP. However, it leaves the N-features of the embedded AgrSP unchecked. This derivation, therefore, is also excluded.

A final possibility would be to move the subject in one swoop from the VP-internal position to the specifier position of the matrix AgrSP, and to insert an expletive *it* in the specifier position of the embedded AgrSP, in order to check the latter's N-features. As this would yield (17b), this derivation must be excluded as well.

To see why (17b) is ungrammatical, consider the status of *it*. Bennis (1986) argues convincingly that expletives of this type are not dummy subjects, but part of the complement domain of the verb *seem* (see also Moro 1993). Therefore, *it* can never be simply inserted in the specifier position of the embedded AgrSP in (17b). If it ends up in that position, it must have raised just like *John*. Apparently, then, *it* can only raise to the matrix AgrSP, and *John* can only raise to the embedded AgrSP. Let us see whether the shortest move requirement is needed to yield this result.

[8] The requirement that syntactic features be eliminated reduces to economy of representation: the interface representations need to have as few symbols as possible. Syntactic features are irrelevant at the interface, and therefore they have to go. Crucially, if they were not eliminated, the interface representations would contain material that is not fully interpretable by the other cognitive systems.

[9] This checking operation was previously called 'Nominative Case assignment'.

Superraising constructions can be pictured schematically as in (19), where both V's are raising verbs:

(19) $[_{AgrSP1}\ [_{VP1}\ V1\ [_{CP}\ [_{AgrSP2}\ [_{VP2}\ V2\ ]]]]]$

John in (17) must originate in the complement of the lower raising verb, V2. *It*, however, can in principle be generated in the direct complement of either V1 or V2. This yields two possible structures for (17):

(20) a. $[_{AgrSP1}$ AgrS1 $[_{VP1}$ V1 [*it* $[_{CP}\ [_{AgrSP2}$ AgrS2 $[_{VP2}$ V2 [*John (to) win*]]]]]]]
b. $[_{AgrSP1}$ AgrS1 $[_{VP1}$ V1 $[_{CP}\ [_{AgrSP2}$ AgrS2 $[_{VP2}$ V2 [*it* [*John (to) win*]]]]]]]

If we start from (20a), (17b) will never be derived. *John* cannot move to the specifier position of AgrSP1, because this would leave the N-features of AgrS2 unchecked. *It* is not available for checking the N-features of AgrS2, because *it* is generated in the matrix clause.

(20b), however, is the structure that might yield (17b), with *it* moving to the specifier position of AgrSP2 and *John* moving to the specifier position of AgrSP1, crossing AgrSP2 in violation of the shortest move requirement. Notice, however, that this derivation would again involve A-movement out of a tensed CP. As is well known, movement out of a tensed CP can only take place via the specifier position of CP, which is excluded in the case of A-movement.

That the 'Tensed CP' constraint is independently needed is clear from the ungrammaticality of examples like (21a), which cannot be explained by the shortest move requirement:

(21) a. * $John_i$ seems [t_i won]
b. It_i seems t_i [John won]

As (21b) shows, the ungrammaticality of (21a) is not due to some requirement on the presence of *that* in raising constructions.[10]

The ban on A-movement out of tensed CPs also accounts for the following cases of superraising:[11]

(22) a. * $John_i$ seems it is likely [that we assured them [t_i to win the race]]
b. * $John_i$ seems [that it was told t_i [that he would win]]

The constructions in (22) have in common that *John* cannot be Case-licensed in the position marked by the trace, and moves to the specifier position of the matrix AgrSP across *it*. However, the raising to the matrix AgrSP in both cases takes place out of a tensed CP, which is not allowed. Again, the shortest move requirement appears to be redundant.

In sum, the superraising constructions in (17) can both be excluded without taking recourse to the shortest move requirement of economy of derivation.

[10] For arguments that the Tensed CP constraint is also needed for A'-movement, see Lee (1993).

[11] These cases were brought to my attention by Phil Branigan and Luigi Rizzi, respectively, as cases that my analysis in Zwart (1993b) leaves unexplained. In Zwart (1993b), I argued that (20b) is an inadmissable basic structure, because *it* cannot be combined with a nontensed clause (which still seems to be a correct generalization).

6. Wh-island facts present a third class of phenomena suggesting an analysis in terms of a shortest move requirement.

(23) *What did he wonder where John put t?

In (23), *what* moves out of an embedded interrogative clause, crossing the position occupied by *where*. This position (the specifier position of the embedded CP) being a potential landing site for *what*, minimality (7) is violated.

The analysis of wh-movement as being *successive cyclic*, first put forward in Chomsky 1973, had a major impact on grammatical theory. It is probably correct to state that the desirability of having short steps in a derivation goes back to this pioneering work.

It is ironic, therefore, that precisely in this domain the superiority of the fewest steps requirement over the shortest move requirement is so apparent. Chomsky 1992:21f, modifying Chomsky 1973:243, explicitly states that successive movement does not consist in a succession of identical adjunction operations but in a single operation *Form Chain*. This operation turns a representation like (24a) into, for instance, (24b), without the intermediate step that yields (24c):

(24) a. e [you think [e [you love who]]]
b. who [you think [t [you love t]]]
c. e [you think [who [you love t]]]

It would take two steps to get from (24a) to (24b) via (24c), whereas *Form Chain* derives (24b) from (24a) in one step.

We may think of *Form Chain* as comprising two independent operations: a movement operation taking the longest step that is needed, and an insertion operation creating intermediate traces where necessary. The movement part of *Form Chain* is in keeping with economy of derivation, viewed as a 'fewest steps' requirement. The insertion part has nothing to do with movement, but is necessary to comply with conditions on chain formation.

As noted in section 4, conditions on chain formation are basically representational in nature. Assuming dependency relations to be local (Koster 1987), chain links must be kept minimal for a chain to be able to receive an interpretation as such. In (23), *what* and its trace are not in one local domain, which makes it difficult to interpret *what* as the displaced object of *put*.

As is well known, and in fact noted in Chomsky 1973:244 note 25, judgments of wh-island violations show considerable variation. Keeping to the simplest case, argument extraction out of wh-islands is much easier than adjunct extraction out of wh-islands. If the shortest move requirement existed, both types of extraction should be equally bad. If the shortest move requirement does not exist, the difference must be explained in different terms.

Work by Cinque 1990 suggests that argument traces may receive an interpretation independently from its dependency relation with the wh-phrase, namely as a small *pro*. This confirms the idea that what is crucial in the judgment of wh-island constructions is the interpretation of each of the elements involved, either as part of a chain, or otherwise.

Ultimately, this suggests that the appearance of traces in constructions like (24b) is due to a principle of *Full Interpretation* rather than to economy. Again, the shortest move requirement seems completely irrelevant.

In fact, we can make an even stronger argument against the shortest move requirement. Let us assume that embedded interrogatives are CPs containing a feature [+wh] (or, as proposed in Müller & Sternefeld 1993, Hoekstra & Zwart 1993, that interrogative clauses are WhPs). This assumption potentially explains effects on (embedded) complementizers in long distance wh-movement constructions, such as the appearance of the wh-complementizer *of* in the following Dutch paradigm (from Hoekstra & Zwart 1993):

(25) a. Piet denkt (*of) dat Jan het gedaan heeft
Pete thinks if that John it done has
'Pete thinks that John did it.'
b. Wat denkt Piet (of) dat Jan gedaan heeft?
what thinks Pete if that John done has
'What does Pete think John did?'

If both the embedded CP and the matrix CP contain a feature [+wh], each position could be the target for the operation *Form Chain*. If economy of derivation contained a shortness requirement, *Form Chain* could only target the embedded specifier position. As this would leave the [+wh] features in the matrix CP unchecked, the interface representation would end up consisting of too many symbols, yielding a crashing derivation. This replicates the conflict between shortest move and fewest steps mentioned by Chomsky 1992:21, which he tried to solve by introducing the notion *Form Chain*. Apparently, the shortness requirement does not apply to *Form Chain*: it is allowed to stretch as far as is necessary. Again, only the number of steps is relevant for economy of derivation.[12]

7. It thus looks like the shortest move requirement can be dispensed with in the analysis of head movement, superraising, and wh-movement. Since these are the core cases giving rise to the shortest move requirement, the optimal hypothesis would seem to be that this requirement is not part of economy of derivation.

This hypothesis finds further support if we consider the process of *Form Chain* in more detail. Recall that chains resulting from long distance wh-movement are created in one step (Chomsky 1992:21). Assuming, as before, that intermediate traces occupy designated positions (the specifier position of CP, or WhP), this raises a question regarding strict cyclicity.

As in Chomsky 1973, Chomsky 1992 assumes that derivations are subject to a condition of strict cyclicity. In Chomsky 1992, this condition is expressed in terms of the bottom-up nature of the representation building process. Each step in a derivation

[12] The German construction in (i) is interesting from this perspective. Here, the wh-phrase is only moved part way, and the specifier of the matrix CP is occupied by a neutral wh-word *was* 'what'. Huybregts 1992 analyses *was* as a quantificational precursor of LF-operator movement. This would imply that partial wh-movement constructions involve two steps, one overt and one covert, in violation of the fewest steps requirement of economy of derivation. I would be forced to assume that in fact the representation in (i) is final, and that no further wh-movement takes place between the Spell Out point (S-structure) and the LF interface point.

(i) Was glaubst du <u>mit wem</u> ich gesprochen habe <u>t</u> ?
what think you with whom I spoken have
'Who do you think I talked to?'

targets a given representation and extends it (Chomsky 1992:32, see also Kitahara 1993). No step can target a position internal to a given representation. Movement is just a process of extracting an element from a given representation and adjoining it, thereby extending the representation (a *singulary operation*). Insertion is a process of extending a given representation by adding another phrase marker (a *binary operation*). [13]

(26) *Strict Cyclicity*
Don't add anything inside a representation

Now if long distance wh-movement is performed by the operation *Form Chain*, this operation will not only extend a given representation, but also stick in WhPs and intermediate traces *inside* this representation. This is a violation of the strict cyclicity condition.[14]

This problem can be solved by analyzing *Form Chain* as consisting of a number of binary operations and a single singulary operation.[15] The binary operations extend an AgrSP to a WhP, first by adjoining an AgrSP containing a wh-element to a functional head Wh, which yields a WhP, and then by adjoining an empty wh-element to this WhP (i.e. in the specifier position of WhP, see Zwart 1993b). The adjunction of the wh-element is necessary to check the N-features of Wh.

These operations, we assume, take place optionally in every clause. If they are left out, the derivation will crash at some point if the clause contains a wh-element. This is because the specifier position of WhP is the designated position for licensing wh-elements.

Suppose the WhP resulting from these operations is then adjoined to a lexical head V, turning the WhP into an embedded interrogative. The V will expand to a full AgrSP, and will then be adjoined to a functional head Wh again. This head Wh again has N-features wanting to be checked. We can do this by adjoining an additional empty wh-element, and then carry on as before. But suppose the WhP is a matrix clause, i.e. not adjoined to a lexical head V in the next step of the derivation. Then the

[13] If Kayne (1993) is correct, this extension can only take place on the left boundary of a representation.

[14] Technically, the *Form Chain* operation does not violate strict cyclicity, because *Form Chain* inserts traces *and* extends the structure at the same time. This, however, depends on the exact formulation of the Strict Cycle condition. In Chomsky 1992, the Strict Cycle Condition is taken to hold for substitution only. As a result, head movement (a subcase of adjunction) never violates strict cyclicity. However, as argued in Zwart 91993b), there is not much reason to distinguish substitution and adjunction in a generalized transformation-based approach to phrase structure (see also Frank and Kroch 1992). This suggests that the apparent exemption of head movement from the Strict Cycle Condition must be derived in a different way. Notice that head movement and XP-movement can be distinguished by the type of mother projection that is created in the adjunction process. In head movement, an X° is created, whereas in XP-movement an XP is created. Violations of the Strict Cycle Condition typically project an XP internal to an XP. In Zwart (1993b) I propose that this is in fact what the Strict Cycle Condition prohibits (assuming no bar level distinction between X' and XP):

(i) *Strict Cyclity*
Do not project an XP inside an XP

Under this formulation, head movement does not violate the Strict Cycle Condition. *Form Chain* however does, because the intermediate empty elements can only be accommodated (under the generalized transformation approach) by projecting an additional WhP node.

[15] The binary operations do not add to the cost of a derivation, since tree structures cannot exist without binary operations.

wh-features of the wh-element contained in the AgrSP we started out with will remain unchecked. Therefore, there is no other option than to apply a singulary operation: moving the wh-element to the matrix WhP in one giant step.

After this, the empty wh-element in the intermediate WhP will start functioning as an intermediate trace, linking the wh-element to its (initial) trace. Given locality conditions on dependency relations, this is the only way to establish a link between the displaced category and its trace.

This derivation of long distance wh-movement (actually, a decomposition of the operation *Form Chain*) is fully compatible with the strict cyclicity condition (26). It contains only one movement operation, in compliance with the fewest steps requirement of economy of derivation. It is in blatant violation of the shortest move requirement of economy of derivation. But this violation is inevitable if we want to make *Form Chain* compatible with the strict cyclicity condition.

In other words, it follows from the strict cyclicity condition that the shortest move requirement is not part of economy of derivation.[16]

8. Let us now return to the Equidistance Principle of Chomsky 1992. This principle modifies the shortest move requirement of economy of derivation. It allows elements to cross a position where they could have landed, provided the target position is in the same minimal domain as the position which is crossed.

(27) *Equidistance* (Chomsky 1992:24):
If α,β are in the same minimal domain, they are equidistant from γ.

As we have seen, positions end up being in the same minimal domain as a function of head movement. This leads to the prediction that objects can only leave the VP, crossing the subject position inside VP, if the verb has moved to AgrO first.[17]

As argued in section 2, this prediction is not borne out. This suggests that the equidistance principle as stated in (27), under the definitions understood, is wrong. But since the equidistance principle is merely a modification of the shortest move requirement, we are not surprised to find that it has no empirical substance. As the shortest move requirement does not exist, a modification of this requirement is bound to make the wrong empirical predictions.

The shortest move requirement, including the equidistance principle, makes one important contribution that is lost in the present approach. Chomsky 1992:26 derives from this requirement the fact that movement of subject and object to their respective licensing positions (*L-related* movement) is always crossing instead of nesting. It is well-known that movement to operator positions (*nonL-related movement*) is nesting rather than crossing (Pesetsky 1982).

This follows in Chomsky's analysis, because (for reasons not discussed here) a minimal domain contains at most two equidistant specifier positions. As a result, if the subject moves to the specifier position of AgrO, so that the subject and its trace occupy the two positions that are equidistant from the object, the object cannot move

[16] As pointed out to me by Anthony Kroch, the strict clycle condition can actually be derived from economy of derivation under the assumptions of this paper. If economy of derivation contains only a 'fewest steps' requirement, it can also be formulated as a 'longest move' requirement. The longest move is always the one that extends the structure, hence adjunction internal to a structure is excluded.

[17] See Chomsky 1992 for the definition of 'minimal domain'. See also Zwart 1993c.

at all. Consequently, the object has to hop over the subject first, and only then is the subject allowed to cross the object, again as a function of head movement creating the proper equidistant landing site.

This important result is lost.[18] We could achieve the same result by stipulating the order of agreement projections in the functional domain, but that would clearly be inferior to Chomsky's approach. This, then, is the problem to solve if we accept the point argued for here, namely that the shortest move requirement does not exist.

9. Accepting the arguments against the existence of the shortest move requirement, economy of derivation can be stated as in (2):

(2) *Economy of derivation*:
Use as few steps as possible in deriving an output representation

The question arises whether economy of derivation can be reduced to economy of representation:

(1) *Economy of representation*:
Use as few symbols as possible in an output representation

Inasmuch as steps reduce the number of symbols, (2) can be reduced to (1).

Movement is generally regarded as a chain creating process. If a chain counts as one symbol, movement will not decrease the number of symbols. However, N-features and V-features must count as symbols, or otherwise they would not be relevant for the principle of Full Interpretation:

(28) *Full Interpretation*:
In an interface representation, do not use useless symbols.

Therefore, only movements that result in checking and elimination of features is allowed by (1).

We can also state that strong features count as symbols at the PF interface, but weak features do not. Movement to check and eliminate strong features then reduces the number of symbols, but movement to check and eliminate weak features does not. *Procrastinate*, then, can be reduced to economy of representation holding for the PF interface.

Greed follows from (1) if we take (1) to apply in a non-global fashion, that is: every element must carry as few symbols as possible in the output representation. Thus, (1)

[18] However, as pointed out to me by Liliane Haegeman and Gereon Müller, the Equidistance Principle as stated is not able to accommodate derivations involving two internal arguments and one external argument. Assuming both internal arguments to have designated licensing positions in the functional domain (i.e. there is a sequence of two AgrOPs), the higher internal argument will at some point in the derivation have to cross the basic position of the external argument (i.e. the Spec,VP) and the licensing position of the lower internal argument (i.e. the specifier of the lower AgrOP) on its way to its licensing position (i.e. the specifier of the higher AgrOP. Spec,VP and the two Spec,AgrOPs, however, cannot all three be equidistant from the basic position of the higher internal argument. The Equidistance Principle appears to be unable to derive these structures. It seems, then, that the Equidistance Principle captures the crossing requirement on A-movement only partly.

forces a feature bearing element α to move, but prohibits movements that do not result in the elimination of a feature carried by α.

If this conception of *Greed* is correct, economy of derivation (i.e. the 'fewest steps' requirement) and the principles that fall under economy of derivation (the *inertness* principles) can be seen to reduce to economy of representation.

Finally, note that economy of representation can be reformulated as in (29):

(29) *Economy of representation (reformulated)*
In an output representation, do not use superfluous symbols.

It is clear from this reformulation that economy of representation, now incorporating economy of derivation, is equivalent to the principle of *Full Interpretation* (28).

10. In conclusion, I have argued that economy of derivation does not contain a requirement that steps be as short as possible. Restrictions on head movement, superraising, and wh-movement follow from well established principles and mechanisms made explicit in recent minimalist work. This resolves the conflict between the 'fewest steps' requirement and the 'shortest move' requirement in Chomsky 1992. I argued that Chomsky's *Form Chain* mechanism must be decomposed into a trace insertion mechanism and a movement mechanism, where the traces are inserted before the movement takes place, in accordance with strict cyclicity. Finally, I offered some speculation on the possibilities of reducing economy of derivation to economy of representation, and economy of representation to the principle of Full Interpretation.

It is interesting to note that, if we were correct in the above, generative grammar, even in the purely derivational approach that incorporates the structure building process of generalized transformations (including movement), does not need any constraints beyond conditions on representation.[19]

References

Bennis, H. (1986) *Gaps and Dummies*. Dordrecht: Foris.

Chomsky, N. (1973) ''Conditions on Transformations''. In S.R. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*, New York: Holt, Rinihart, and Winston.

Chomsky, N. (1986) *Barriers*. Cambridge: MIT Press.

Chomsky, N. (1991) ''Some Notes on Economy of Derivation and Representation''. In R. Freidin, ed., *Principles and Parameters in Comparative Grammar*, Cambridge: MIT Press.

Chomsky, N. (1992) ''A Minimalist Program for Linguistic Theory''. *MIT Occasional Papers in Linguistics* 1. Cambridge: MIT.

Cinque, G. (1990) *Types of A'-Dependencies*. Cambridge: MIT Press.

[19] This final remark presupposes that bounding conditions (yielding for instance the 'tensed CP' constraint) and conditions that ensure the correct order of projections in the functional domain are to be considered as representational rather than derivational conditions. As for the conditions on interpretation requiring a moved element to be related to its trace, it was assumed in the text that these are either representational conditions, or extragrammatical conditions of interpretation.

Frank, R. and A.S. Kroch (1992) ''Generalized Transformations in Successive Cyclic and Long Dependencies (or Substitution + Extending = Adjoining)''. Paper presented at the GLOW Colloquium, Lund, April 6.

Hoekstra, E. & C.J.W. Zwart (1993) ''De structuur van de CP: Functionele projecties voor topics en vraagwoorden in het Nederlands''. To appear in *Spektator*.

Huybregts, R. (1992) ''Scattering: A Derivational Approach without S-structure''. Paper presented at the 8th Workshop on Comparative Germanic Syntax, Tromsø, November 21.

Kayne, R.S. (1993) ''The Anti-Symmetry of Syntax''. Ms., City University of New York.

Kitahara, H. (1993) ''Target α: Deducing the Strict Cycle Condition from Principles of Economy''. Paper presented at the GLOW Colloquium, Lund, April 6.

Koster, J. (1987) *Domains and Dynasties. The Radical Autonomy of Syntax*. Dordrecht: Foris.

Lee, R.K. (1993) ''Relative Minimality and Tensed Island as Two Separate Derivational Constraints''. In V. Lindblad and M. Gamon, eds., *Papers from the Fifth Student Conference in Linguistics (MIT Working Papers in Linguistics 20)*, 145-160.

Moro, A. (1993) *I predicati nominali e la struttura della frase*. Padova: Unipress.

Müller, G. & W. Sternefeld (1992) ''Improper Movement and Unambiguous Binding''. *Linguistic Inquiry* 24, 461-507.

Ouhalla, J. (1989) ''Against the Head Movement Constraint.'' *UCL Working Papers in Linguistics 1*, 193-205.

Pesetsky, D.M. (1982) *Paths and Categories*. MIT: PhD dissertation.

Rizzi, L. (1990) *Relativized Minimality*. Cambridge: MIT Press.

Travis, L. DeMena (1984) *Parameters and Effects of Word Order Variation*. MIT: PhD dissertation.

Vikner, S. (1991) *Verb Movement and the Licensing of NP-Positions in the Germanic Languages*. University of Geneva: PhD dissertation.

Wilder, Chr. and D. Ćavar (1993) ''Word Order Variation, Verb Movement, and Economy Principles''. Ms., University of Frankfurt on Main.

Zwart, C.J.W. (1992a) ''Matching''. In D. Gilbers and S. Looyenga, eds., *Language and Cognition 2. Yearbook 1992 of the Research Group for Linguistic Theory and Knowledge Representation of the University of Groningen*. Groningen: University of Groningen.

Zwart, C.J.W. (1992b) ''SOV Languages are Head Initial''. Paper presented at the 8th Comparative Germanic Syntax Workshop, Tromsø, November 21. To appear in *The Linguistic Review* as ''Dutch Is Head Initial''.

Zwart, C.J.W. (1993a) ''Verb Movement and Complementizer Agreement''. In J.D. Bobaljik and C. Phillips, eds., *Papers on Case and Agreement I (MIT Working Papers in Linguistics 18)*, 297-340.

Zwart, C.J.W. (1993b) *Dutch Syntax: A Minimalist Approach*. University of Groningen: PhD dissertation.